PHD

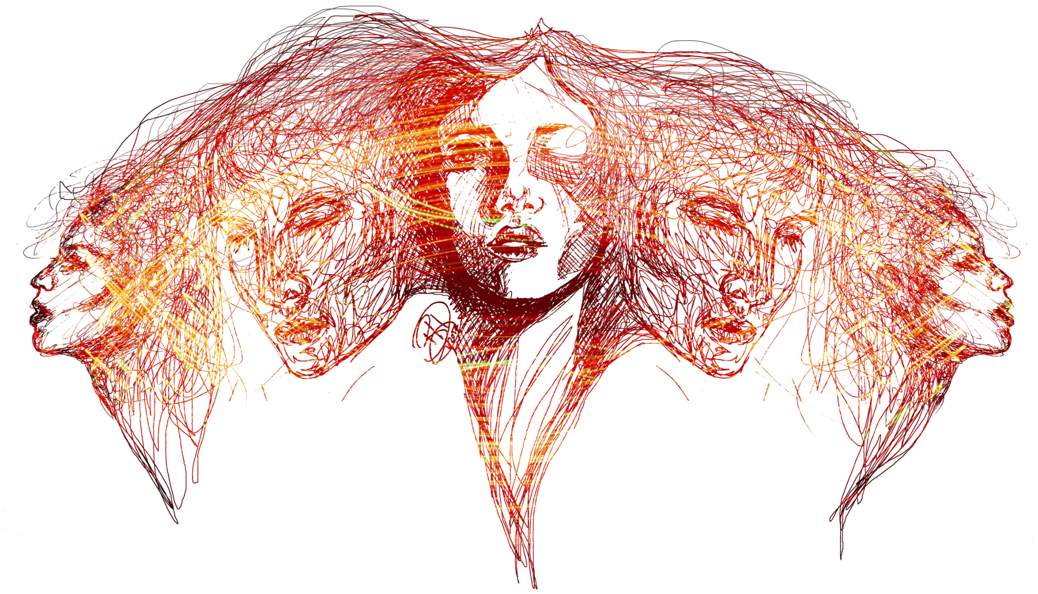
Shape Shifting Behavior and Identity Across Digital Systems

Davidson, Brittany

Award date:
2019

Awarding institution:
University of Bath

Shape Shifting Behavior and Identity Across Digital Systems



Brittany Isabella Davidson

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
School of Management
Information, Decisions and Operations Group

27th June 2019

ACKNOWLEDGEMENTS

This thesis is lovingly dedicated to the memory of my mother, Alexandra R. Davidson, who passed away just before I started this journey, on the *27th June 2016*. The first submission of this thesis marks three years without you. Her sheer determination, courage, and spirit that she willfully applied to every situation she found herself in inspired me to keep going and find that in myself when I became lost and not knowing where I was heading. While this thesis certainly would not have been your favorite topic to read, this is for you, mom.

I wouldn't be here without the endless support and unconditional love from my father, Philip R. Davidson, whom I also cannot find enough words to thank for standing by me throughout this process. Dad, this thesis is equally for you, thank you for pushing me to keep going when times were good, and especially when I had doubts.

David, thank you for your endless kindness and patience with me – for being my support network, for reading draft after draft of many chapters of this thesis, listening continually, and perhaps most importantly, for believing in me when I didn't. Additionally, my knowledge and work has developed tremendously from working with you; thank you for challenging me and continuing to do so.

Ollie, Jo, and Mehrnaz – it's been the most tremendous pleasure, from the whisky adventures, to the long summer evenings in Crystal Palace, the laughter, and the frantic stressing (almost always from me) – you both continually picked me up. Thank you for being there always. Fluffy, Badger, Tealeaf, and Strawberry…you know who you are. Thank you for the endless reassurance and love starting at the University of Edinburgh, to Bath, and onwards to Bristol. Similarly, Heather – thank you for your kindness and support over the final stages of my PhD – I only hope I can return the favor!

I owe a huge thanks to my supervisor, Professor Adam Joinson, for challenging me and affording me so many opportunities during my PhD. Thank you for your tremendous support and compassion. Lukasz, thank you for your kindness. As always – Sooz, thank you for looking after me throughout this process. You have been wonderful, and I cannot thank you enough for your support.

Finally, Professor Mark Levine and Professor Brian Squire – thank you for taking the time to examine my thesis and provide such valuable insights and discussion for my work moving forward.

* * *

I would like to wish some final thanks to for the funding bodies that allowed this work to happen. First, thank you to everyone within the Centre for Research and Evidence on Security Threats (CREST) for so much support; from conference costs to further studies both abroad in the University of Amsterdam and locally at the University of Essex, as well as supporting my time working in the Cabinet Office during my time with Bath.

Additionally, thank you to the University of Bath, IDO Division in the School of Management for awarding me a URS scholarship to support my time here.

Contents Page

List of Figures

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

Digital devices and services have become ubiquitous and persuasive. Hence, it is critically important to understand these digital interactions (or digital traces) as daily life involves rapid transitions between various offline and online interactions.

In this thesis, I explored how individuals change and adapt their behavior across a variety of digital systems through the lens of *social role* and *identity theory* perspectives *computationally*. Individuals naturally adapt their behavior across contexts. This social flexibility is integral to human interaction. For instance, we would not expect someone to behave identically at work and home. Here, I extend this concept to the digital world, where behavior and identity is examined across a variety of digital devices and systems utilizing new methods and approaches.

I first made the case for using new methods to better understand digital behaviors by demonstrating the discrepancy between objective technology use and self-reported behavior, which provides an example of an alternative method for psychologists. Next, I considered the dynamic nature of social roles online, for the first time, by utilizing supervised and unsupervised machine learning techniques. The results show how roles relate to leadership online, and which pathways users took to become leaders. I then analyzed user linguistic flexibility via computational methods (linguistic style matching (LSM)) across different contexts. This revealed that users typically diverged their linguistic style from the community. This happened to a lesser extreme within strictly moderated contexts. Finally, I undertook a qualitative approach to user experiences across online systems, which further showed how users negotiate their self-presentation and identity online.

Overall, this thesis confirms that new, data intensive, objective measurements can sit alongside traditional approaches. If these are adopted more widely, social psychology will move further beyond the lab when it comes to understanding how people live across digital and real-world environments.

CHAPTER I

INTRODUCTION

1 General Introduction

A life without (digital) technologies no longer exists – from the printing press, the telephone, to television, film (analog and digital), computers, the internet, smartphones, and beyond. The mass-adoption of (communication) technologies has transformed how we live. Technology is inherently there to make our lives 'easier': we are assisted by short, rapid communication via various forms of messaging (e.g., iMessage, SMS, WhatsApp) and social media (e.g., Facebook, Twitter); our home-based errands are made easier (e.g., cooking, cleaning); our workplaces transformed (e.g., email, business analytics); and we might even be 'nudged' to live healthier and longer (e.g., via wearables, or other devices to help monitor conditions) (e.g., Ellis, 2019; Kuijer & Giaccardi, 2018; MacKenzie & Wajcman, 1999; Mitzner et al., 2010).

This integration of technology is certainly not a new phenomenon. From the early use of tools in the 'Metal Ages' – which led to rapid and far reaching social changes (e.g., Gilman et al., 1981) – to the development of the printing press in fourteenth century in Germany (Dittmar, 2011), technology has transformed society for centuries. More recently we have seen developments such as artificial intelligence (AI), the internet of things (IoT), virtual or augmented reality (VR/AR), self-driving cars, robotics and smart cities, that promise further deep impacts on society. As these technologies become more persuasive and ubiquitous, understanding the impacts of these technologies on individual and group behavior is critically important (e.g., MacKenzie & Wajcman, 1999).

It is therefore of no surprise that technology usage, technology integration into everyday life, technology adoption, or even resistance to technologies, are topics of keen interest across social science research. This work, often under the guise of *Science and Technology Studies* (STS), has taken a range of approaches, usually grounded in sociological theories, to understand not only how a technology impacts individuals and society, but also how technology is inherently economic, political, and social (KacKenzie & Wajcman, 1999; Mackay & Gillespie, 1992). Critically, STS has tended to eschew the view that technology is neutral, instead viewing the interaction between society and technology as one that has both positive and negative impacts on human culture (e.g., Geels, 2007). This extends beyond simply the use of technology with ill-intent or unethically (MacKenzie & Wajcman, 1999; Sawyer, 2019). For instance, current debates around 'big data' and AI are increasingly focused on the ways in which the fields of ethics, law, and public policy can protect society in a new digital age (Wachter & Mittelstadt, 2018).

I would argue that a deeper knowledge and understanding of how individuals and groups use various technologies can aid the development of better technologies, and that increased knowledge of the impacts of technologies on individuals, groups, and society will help us to shape the ways those technologies are implemented and thus the impact they have. Further, understanding what the future of technologies may be across disciplines (e.g., precision health/medicine, defense and security, workplace enhancements) can help place social concerns at the heart of technological innovation. We should also consider what these new technologies can also offer social psychological science research specifically (see section 1.3).

Technology continues to develop and change rapidly (Manogaran, Thota, & Lopez, 2018; Roser & Ritchie, 2019). Alongside these innovations, there has also been a sudden influx of new information and data at an unprecedented scale. Every single interaction someone has online or on their digital device leaves behind digital traces or digital footprint (Weaver & Gahegan, 2007). It is estimated that there will be 75 billion Internet of Things (IoT) devices connected by 2025 (Columbus, 2016). There are 6,940 matches on Tinder every minute, alongside 13 million texts sent, and 4 million google searches (Ahmad, 2018) – each providing its own digital trace about its users. User digital traces continue to grow (Latour, 2007; Manogaran, Thota, & Lopez, 2018), not least due to people's frequent and habitual use of technologies (e.g., Andrews, Ellis, Shaw, & Piwek, 2015; Ellis, Davidson, Shaw, & Geyer, 2018). This digital trace data provides an abundance of opportunities across research disciplines (Manogaran, Thota, & Lopez, 2018) – from aiding 'basic' research that seeks to understand how and why we use technology to studies that seek to predict technology use, draw implications of technology use, and or develop our theoretical understanding of technology and society. The goal of the research in the present thesis is to contribute to this body of work, both by seeking to develop and apply new methodological approaches for the study of people's behavior via technology and by gaining new insights about how people behave when interacting using new technology.

1.1 Thesis Context and Research Questions

Humans are fundamentally social beings, something that has continued as the world has become increasingly digitized (e.g., Dellarocas, 2003; Loebbecke & Picot, 2015). Technologies have arguably improved and provided opportunities across daily life, for example, transport, entertainment, medicine, building development (housing or work-related), and communication (MacKenzie & Wajman, 1999). In many ways, various technological developments have facilitated our social networks, both on- and offline. For instance, the telephone and various forms of transport drastically changed our networks by allowing increased contact with others by aiding the ability to travel further, but to also speak with others without needing to be in close proximity. This digitization of communication across devices and systems allows us to meet people across all parts of the world (e.g., in online communities, dating apps, email, social networking sites) (Davidson, Joinson, & Jones, 2018). Of course, the very nature of our communication has also changed, with almost instant communications across formats, from text, voice, to video – from broadcasting across mass-communication sites (e.g., Twitter and Facebook), to smaller 'group chats' (e.g., WhatsApp) (Boulianne, 2018). This variety of communication channels (e.g., broadcasting versus a direct message to someone) impacts how we interact as naturally the audience and context changes, similar to our offline behavior (seen in *Chapters IV* and *V*). People will naturally change and adapt across settings, where one is likely to have a different pattern of behavior in the workplace compared to when interacting with family and friends (Fiske, 2010). This social flexibility is natural and expected as we transition from home to meeting friends, and we could anticipate that similar contextual changes in behavior may also be seen online; as a user moved from posting photos from their holiday to Facebook and Instagram, to updating their LinkedIn page.

It is clear that various technologies have become highly integrated into our everyday lives (e.g., Shaw et al., 2018), which provides an abundance of research avenues to follow. The research in the present thesis looks at people's interactions with digital technologies, specifically, *how* they behave across different systems, for example, online communities (e.g., Reddit), social networking sites (SNS) (e.g., Facebook, Instagram), as well as across different devices (e.g., smartphones, laptops). Further, this research will consider several different methodologies to understand user behavior across systems and devices (e.g., correlational studies, machine learning techniques, repertory grid technique, text analysis, etc.). Additionally, there is a question regarding whether the data, methods, analysis, and inferences made within research focused on technology and society and is indeed appropriate for the questions at hand. Particularly towards the end of *Chapter I* and *Chapter II* will address this issue.

Hence, the overall research questions are addressed by the research in this thesis are;

1. Do individuals adapt their behavior across different systems or over time? If so, how can this be measured and theorized?
2. What can we understand about an individual's interactions with technologies from a variety of approaches, data sources (e.g., usage, meta-data), and methods?

1.2 Behavior Across Digital Systems

A fundamental part of human sociality is our ability to change and adapt our behavior in response to both internal (e.g., emotion, expectations) and external (e.g., audience, environment) factors (e.g., Herrmann, Jahnke, & Loser, 2004; Hogg, Terry, & White, 1995). This social flexibility extends to our behavior across contexts – we wouldn't expect a person to behave identically as s/he transits from work to home, or from home to a night out with friends. This type of context-based behavioral change is typical and expected by those around us – and has been theorized as *social role theory* (e.g., Fiske, 2010; Herrmann et al., 2004).

It would be expected that the same variance in social roles and behaviour that are observed in the offline world would be also be seen in online settings. Indeed, social media for most is a diverse experience (*Chapter IV* and *V*). For example, a colleague may be inclined to present themselves differently on LinkedIn compared to on Facebook, while their Tinder or Reddit profile may be almost unrecognizable in comparison. As such, we anticipate that users will behave differently as they move from one online system to another. Indeed, there is some limited evidence for this – for instance, Vasalou and Joinson (2009) found that people were likely to create a more attractive avatar for a dating profile and a more 'intellectual looking' avatar for an online gaming profile.

This does not mean that online identity is in some way not 'authentic' or 'genuine', but rather that identity is itself dynamic and changing – aligning with the notion of *social roles*, where different elements of multiple identities are drawn into use according to the context (e.g., social norms, audience(s)), ongoing interaction, and also the way the system was intrinsically designed (Levina & Arriaga, 2014). However, while individuals may be able to manage audiences and maintain a separate work-life balance, this is becoming increasingly difficult online, especially in an age where social media platforms have become ever more interlinked and ubiquitous.

Increasingly, services like WhatsApp and Instagram are now parts of a single organization: Facebook (Facebook, 2014). The acquisition of these services often leads to shared authentication and access routes, merged content, and contacts suggested from one platform to the next. This of course not only raises security concerns, but also ethical concerns of maintaining privacy across once separate systems. This unremitting data-sharing has in fact led to regulatory warnings in France due to potential violations of French and EU data protection and privacy laws (e.g., Fioretti, 2017). While this merging of multiple sites (with potentially different audiences) might not seem problematic in a world of a '*single, authentic identity*' envisioned by Facebook's CEO Zuckerberg, there is considerable evidence that a 'single identity' is neither natural nor usual in offline life.

There has been extensive research into the 'multiple audience problem' posed by social media (e.g., Binder, Howes, & Sutcliffe, 2009; Marder, Joinson, & Shankar, 2012; Marwick & boyd, 2011). Studies of individual social media sites (usually Facebook) confirm that users often have friends, colleagues, and potentially bosses and parents or wider family connected to their profile as 'friends'. This can be problematic as it can cause anxiety and discomfort due to the discrepancies in terms of audience expectations about who we are, and how we should behave (e.g., Marder et al., 2012; Rui & Stefanone, 2013; van Dijck, 2013). For instance, one might be concerned about their parents or boss seeing photos from last night on Facebook. Similarly, one might not want their Tinder date to see their embarrassing photos from high school.

Hence, the presence of multiple audiences across social media sites means that users may need to actively monitor their self-presentation in order to meet the different expectations of diverse groups (Marder et al., 2012). Presenting multiple facets of ourselves is not well supported on most single services, in part because 'grouping' systems remain under-utilized, for example, restricting the audience for particular posts, albums, or photos on Facebook, or hiding elements of one's profile from parts of a 'friend' list. As a result, many users report presenting to the 'lowest common denominator' where the most easily offended audience acts to 'chill' expression (Marder, Joinson, Shankar, & Houghton, 2016). This causes online audiences to act as a type of information control or management (Hogan, 2010). Boyd (2007) described this as a 'context collapse', where an individual's self-presentations would typically vary due to multiple audiences, but cannot due to a single context (e.g., Facebook). This therefore causes individuals to only share content and information that is deemed acceptable to the broadest audience within their network. For instance, an individual may have a seemingly neutral Facebook profile with little information regarding their sexuality and sexual preferences as they are 'friends' with their boss, family, and socially distant colleagues. In contrast, this same individual might be much more open about this on an online dating profile or other (anonymous or not) online communities.

However, less is known about how users *actually* behave and negotiate multiple identities across different systems and devices. Is the assumption that people separate audiences by using different services correct? In addition, it is not known what the challenges of unifying systems are (e.g., WhatsApp, Instagram, and Facebook) for users. For instance, these systems share more functionality than before (e.g., 'stories' and 'moments') – although, it is not known whether this serves to 'unify' an individual's identity, or if discrete social roles are retained. While functionality and system design certainly impact the individual, these systems are still distinct, meaning that a 'single, authentic identity' will remain a far cry from reality. This can be seen as positive for users because discrete online systems might allow for deeper, more varied self-expression, which ultimately allows users to 'play' and try different identities without the judgement or ridicule of others (Joinson, 2001; Joinson & Paine, 2007). This could be helpful for people who do not have an offline space to speak openly about personal preferences (e.g., sexuality) (e.g., McKenna & Bargh, 1998; Papacharissi, 2002), for those who are exploring or testing new identities, or for people who would wish to relinquish a role when 'off duty' (e.g. teachers, police). Therefore, the research presented in this thesis anticipates that there will be variance in behavior across different systems (e.g., Facebook, Reddit, Instagram, etc.). Additionally, I also investigate how user roles and behavior changes over time within the same community.

1.3 Traditional Methodological Approaches in Psychology

One fundamental aim of psychology is to understand human behavior. In recent years new technologies have created novel opportunities that have not only transformed data collection and analysis, but also opened entirely new avenues for conceptualizing the measurement of behavior. There are numerous ways in which psychological research can be conducted, with experimental approaches (Alcock & Sadava, 2014) and cross-sectional designs (e.g., survey-based studies) often used, alongside a range of qualitative approaches (Bryman, 2008). Experimental designs have had tremendous impact across psychological science, which has laid foundations that are still prominent today. For instance, within cognitive psychology, much of the decision-making and behavioral economics, and heuristics research was experimental (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1973, 1981) and social psychology has often utilized experiments to examine, for example, bystander effects (e.g., Levine, Cassidy, & Brazier, 2002; Levine & Crowther, 2008), conformity (e.g., Asch, 1956), or cognitive dissonance (e.g., Festinger & Carlsmith, 1959).

However, more recently, there is a question regarding the appropriateness and relevance of some traditional approaches in an increasingly technological world for *certain* research questions. For instance, attempting to examine smartphone or social media in a lab may be less useful as there are likely to be discrepancies in behavior in lab-based settings verses at home or online (e.g., Joinson, 1999). For example, a recent experimental study by Kushlev et al. (2019) reported that, '*smartphones reduce smiles between strangers*'. This study consisted of pairs of strangers sitting in a waiting room for 10 minutes with their smartphones (experimental condition) or without their smartphones (control group). As the title suggests, they found that those with their phones smiled *less*. Immediately, this caused concern and was covered across various media outlets (e.g., Dolan, 2018; McDonald, 2018), aligning to a belief that technology, and specifically, smartphones, are inherently bad (Davidson & Ellis, 2019). However, one must ask as to whether this experiment is actually about smartphones at all. If the participants in the experimental condition were given an iPad, book, or there was a television present, the results are perhaps likely to be the same. Hence, one would assume this study is more about whether there is a tangible item or distraction present, people are likely to smile less due to their attention being diverted to <*insert item or device of choice*>, rather than generating knowledge about the technology itself. The fundamental point here is the question as to whether some of these experiments are measuring interactions and behavior with technologies or are these technologies merely a tool or stimulus in the experimental research that is actually measuring something entirely different?

Similarly, there will often be a question over the ecological validity of experimental work, which highlights the discrepancies between participant behavior in controlled laboratory settings and real-world behavior. Arguably, there will be a role dynamic between the experimenter and participant, which may impact the participant's behavior from the outset (Orne, 1962) – also known as the Hawthorne Effect (e.g., McCarney et.al., 2007). The tension between those who wish to maintain ecological validity and those focused on experimental controls and the removal of extraneous variables continues to fuel a long-standing debate (Parsons, 2015). With this in mind, one must be careful in terms of experimental designs and understanding technology use, as the measurement of this use is not necessarily straightforward.

Returning to this thesis specifically, the main research topic is *behavior* across different systems and devices. The overall research question relates to behavior, which calls for measurement of behavior itself with a technology. Typically, the gold standard of technology use measurement in psychological science has been self-report scales or estimates (Ellis, 2019; Ellis et al., 2018). For a variety of reasons discussed throughout *Chapter II*, it is questionable as to whether this is an appropriate measure of usage. With technologies becoming increasingly integrated into everyday life, users become less able to accurately report their usage (Ellis, 2019; Ellis et al., 2018; Shaw et al., 2018). In contrast, many studies also utilize surveys in order to assess experiences, attitudes, and emotion towards various technologies (e.g., social media, smartphones) (e.g., Gangadharbatla, 2008; Marder et al., 2016; Marder et al., 2017), which is as important as understanding the actual behavior and interactions involved with technologies. However, it must be clear when behavior is and is not measured, which is not always the case with much of the recent research often relating to technology and the negative impacts of it on the individual and society (Ellis, 2019; Ellis et al., 2018).

When the research topic of interest is usage, other disciplines (e.g., computer science) have typically measured actual objective usage rather than asking people about their usage (Ellis, 2019) as this provides accurate and reliable insight into what individuals are actually doing with various technologies (e.g., usage time, which apps they are using, location, etc.). Similarly, website scraping techniques can also offer other insights via meta-data or content data (e.g., how people behaved, or what they posted, in the community of choice). Hence, moving forward, this thesis utilizes real-world, objective behavioral data, also known as digital traces (Weaver & Gahegan, 2007). This is, first, in order to have greater confidence in the data, analysis, and inferences made about behavior, as this is explicitly measured. Of course, any measurement will have limitations, which are discussed at the end of each chapter. Second, utilizing different types of data will reveal different insights about the users and their interactions with technology. Hence, each chapter in this thesis will use a different data source and a variety of methods (e.g., objective behavioral (duration) data (*Chapter II*), meta-data (*Chapter III*), content data (*Chapter IV*), and interview data – individual experiences with technology (*Chapter V*)).

1.3.1 Data Overload

Similar to previous 'analog' or traditional research regarding human behavior, there are many approaches, conceptualizations, methods, and analytical techniques that can be employed, and the movement towards digital methods is no exception to this rule. In fact, methods and analytical techniques have had to develop and adapt in order to handle the sheer scale and variety of data that are being generated (Dufresne & Davidson, 2019). This means that analytical techniques have had to shift towards data- and computationally-intensive methods with large samples, noisy, 'real-world' behavior, and often a heavy component of data cleaning and preparation prior to any form of analysis (Dufresne & Davidson, 2019) (*Chapters III and IV*).

While there is a wealth of data at our fingertips, we must not be seduced by this, as not all data are created equal. It can be unreliable, inaccurate, and therefore produce misleading findings and implications, which is arguably more dangerous than having no data to begin with (Alcock & Sadava, 2014). While a challenge of 'big data analytics' is to reuse data for new insights rather than its original purpose, understanding what the data can genuinely offer is critically important (Sawyer, 2019). Data is not always capturing the intended construct, for instance, there is a clear misalignment between self-reported technology usage and objective technology usage (*Chapter II*). This is critically important, as it remains unclear what these self-reported scales are indeed measuring, which creates an issue with inferences and implications of research (Ellis, 2019; Ellis et al, 2018; Andrews et al., 2015). However, this research seeks to provide (new) methods and approaches to understanding technology interaction based on both objective and interview data.

1.3.2 Alternative Approaches and Methods

Naturally, there are many limitations and critiques of experimental, cross-sectional, and other quantitative approaches and methods more generally. Often these relate to the issues with experiments most often occurring in controlled lab settings that do not reflect the 'real world' and 'real world behavior'. Similarly, these approaches by nature also restrict external influences and often treat individual variance often as noise (Alcock & Sadava, 2014). However, there is an increasing shift towards 'behavioral analytics' or computationally intensive methods, which aims to harness and create insights from a variety of datasets. While these methods and technologies are often used in computer science (Ellis et al, 2018; Jones et al., 2015; Oliver, 2010; Zhao et al., 2016), they have much to offer psychological science in terms of behavior across systems and devices based on objective behavioral data (which will be demonstrated in *Chapters II, III,* and *IV*).

These new methods are increasingly data-intensive and computationally-driven, which creates further opportunities. Across this thesis, the fundamental aim has been to analyze both individual and group level behavior. This meant utilizing a variety of methods to transform the granularity of the data depending on the research question. For instance, *Chapter IV* considered the language use of redditors (users of the discussion site 'reddit'), which at the outset involved pre-processing of data and examining posting frequency and content of ~3.1 million observations (see *Chapter IV, Appendix 1*). However, part of the subsequent analysis reduced the original data set further and considered data from 99 individuals who met specific criteria.

These computational approaches can also be used to reduce data dimensionality, as seen in *Chapter III*, which aimed to understand group level behavior. Here we used clustering algorithms to group users based on various behavioral metrics. These are only two examples of the range of analysis that can be done computationally to manipulate and transform the granularity of data. This offers insights across many areas of research as individual differences are an important factor to consider rather than reduce – especially as technology continues to point towards precision (e.g., precision health/medicine, highly tailored advertising across systems, and increased personalization online).

In contrast to newer computational or more traditional quantitative methods, there are alternative approaches that tend to share a philosophical grounding in social constructivism that rejects quantitative perspectives (Alcock & Sadava, 2014). For example, 'Critical Psychology' and 'Discursive Psychology', where critical psychology rejects many notions taught within psychological science and is interested in power, social structure, organization, and institutions (Fox, Prilleltensky, & Austin, 2009), and discursive psychology relates to the action of speech, talking, and writing (Edwards & Potter, 1992).

These qualitative approaches offer much in the way of in-depth understanding of an individual's experiences, attitudes, feelings, and emotions towards something – for example in *Chapter V*, we examined individual experiences with various online systems (e.g., Facebook, Instagram, LinkedIn, Twitter). This is to demonstrate that while most of this research advocates objective behavioral measurement and new approaches, this is not to discount the importance of understanding individual's experiences with technologies as this is equally as important to knowing what individuals objectively do with technologies.

1.4 Approaches and Methods Utilized in this Thesis

The aim of this thesis is to investigate whether identity, self-presentation, and online behavior may change and adapt in online systems, both at the individual and group level. Secondly, the research intends to demonstrate various methods and approaches to understanding online behavior.

This thesis draws upon both primary and secondary data sources, utilizing a variety of approaches and methods in order to analyze these datasets. The following concepts were investigated:

1. Critical validation of traditional methods to measure online behavior (smartphone usage) [Chapter II];
2. Social role changes over time in online communities [Chapter III];
3. Linguistic style across online communities [Chapter IV];
4. Qualitative insights into user's shape shifting identities online [Chapter V];

Table 1. Methodologies used in Empirical Chapters

	Chapter II	**Chapter III**	**Chapter IV**	**Chapter V**
Design	Correlational; between survey and objective data	Longitudinal analysis with meta-data	Linguistic style analyses on content data	Qualitative; semi-structured interviews
Data Source	Surveys; smartphone usage data	Two online communities; scraped data	Scraped reddit data	Primary data; interviews
Sample Size	238	>1,000	>24,000	22
Analysis Type	Correlational, cluster analysis, group comparisons ANOVA	Cluster analysis (*k*-means), classification (naïve Bayes)	Linguistic measures (LIWC), natural language processing (NLP), within-between ANOVA	Thematic analysis, repertory grids, Euclidean distances
Publication Stage*	*Published: IJHCS*	*Published: PLoS ONE*	*Under review upon PhD completion*	*Submitted upon PhD completion*

* Full references are provided in Chapter Overviews below.

1.4.1 Overview of Chapters

This thesis takes adopts the 'alternative format', with each subsequent chapter (until the conclusion) forming a discrete paper. A summary of each chapter is below with additional details of their publication status.

Chapter I

The first chapter of this thesis focuses on the various approaches one can utilize while researching online behavior, the theoretical background of the work presented, and the outline of the rest of the thesis. This chapter provides a critical overview of the various approaches to online behavior, how this translates into more recent approaches to behavior online and the common methods used. The limitations of various approaches are highlighted before providing the steps this thesis takes in order to mitigate them. Finally, it provides an outline of this thesis.

Chapter II

Chapter II has two key purposes. First, it provides a critical insight into the measurement issues seen with technology use in psychology. Second, it attempted to validate several well-known scales used to measure technology use.

Psychology is naturally interested in understanding actual, observable behavior, as well as, thoughts, feelings, attitudes, or ideologies (Alcock & Sadava, 2014). However, when investigating technology usage and interaction, there have been questions regarding the measurement of technology use and whether the methods employed actually predict or correlate with actual behavior of technology use. This is examined and tested within this chapter. This was published in the *International Journal of Human-Computer Studies* as: Ellis, D. A., **Davidson, B. I.**, Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies*, 130, 86-92, DOI: 10.1016/j.ijhcs.2019.05.004.

Chapter III

Chapter III only utilizes objective behavior. Here, meta-data was captured from two online communities that contains the interactions users had made with the site if they had an account. This includes; number of posts, average word counts, network features (in- and out- degree), where they post (e.g., threads or subforums), etc.

The aim of this chapter is to extend literature regarding roles of users at the group level. Much has been written about roles online (e.g., Ang & Zaphiris, 2010; Pfeil, Svangstu, Ang, & Zaphiris, 2011; Welser et al., 2011; Welser, Gleave, Barash, Smith, & Meckes, 2009), however, these tend to only examine roles at a single point in time. Hence, in the present chapter, we reveal the roles present in two online forums, and we examine role changes over a two-year period of one community. We demonstrate that users can, and do, change roles over time within communities, and present maps of the pathways users most commonly take when changing roles. We ground this work using one of the most influential theories of engagement and leadership within Human-Computer Interaction: The Reader-to-Leader Framework (Preece & Schneiderman, 2009). This paper has been published in *PLOS ONE* as: **Davidson, B.**

I., Jones, S. L., Joinson, A. J., & Hinds, J. (2019), The evolution of online ideological communities, PLoS ONE, 14(5): e0216932, DOI: 10.1371/journal.pone.0216932.

Chapter IV

Similar to *Chapter III, Chapter IV* also only utilizes objective behavior. In this case, content data from Reddit has been analyzed at a large scale. The premise of this paper is to understand how user behavior, specifically linguistic style, may or may not change across online subreddits, which one can argue are different communities.

The dataset was sourced via Google datasets (Syed, Voelske, Potthast, & Stein, 2018). This chapter analyzes function words and user's linguistic style match (LSM) to specific communities. We compared sets of users who had all posted in the same subreddits to examine whether their linguistic style converged or diverged from various subreddits using Communication Accommodation Theory (CAT). Additionally, this chapter tested whether explicit differences in moderation impacted user linguistic style matching with a community.

Chapter V

Chapter V, in contrast to the previous chapters, takes an alternative approach. Here, traditional qualitative approaches and methods are employed in order to understand online behavior. This paper aimed to explore how and why individuals will change and adapt their behavior online, from their self-presentation, behavior (e.g., content they share online), and the audience (e.g., professional vs non-professional) they have for each social media site (e.g., Facebook, Instagram, Snapchat, etc.). Semi-structured interviews were conducted, which was analyzed using Braun and Clarke's Thematic Analysis (2006). Additionally, Repertory Grids were also used in this chapter to further reveal similarities and differences in experience with social media sites.

Chapter VI

Chapter VI is the final 'chapter' of this thesis. It simply provides an overview of the findings from each chapter based on the two research questions put forward in *Chapter I*. It provides an additional overview of future research directions and final concluding thoughts.

References

Ahmad, I., (2018). How Much Data Is Generated Every Minute? [Infographic]. *Social Media Today*. [Online]: https://www.socialmediatoday.com/news/how-much-data-is-generated-every-minute-infographic-1/525692/. [Accessed on]: 5/9/2019.

Alcock, J., & Sadava, S. (2014). *An Introduction to Social Psychology: Global Perspectives*. (M. Carmichael, Ed.). Los Angeles | London | New Delhi | Singapore | Washington DC: Sage.

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS ONE*, *10*(10), 1–9. https://doi.org/10.1371/journal.pone.0139004

Ang, C. S., & Zaphiris, P. (2010). Social Roles of Players in Mmorpg Guilds. *Information, Communication & Society*, *13*(4), 592–614. https://doi.org/10.1080/13691180903266952

Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, *70*(9), 1–70.

Binder, J., Howes, A., & Sutcliffe, A. (2009). The problem of conflicting social spheres: effects of network structure on experienced tension in social network sites. *In Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 965–974).

Boulianne, S., (2018). Twenty Years of Digital Media Effects on Civic and Political Participation. *Communication Research*. https://doi.org/10.1177/0093650218808186

boyd, danah, & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*, 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bryman, A. (2008). *Social Research Methods* (Third Edit). Oxford, UK: Oxford University Press.

Columbus, L., (2016). Roundup of Internet Of Things Forecasts And Market Estimates, 2016. *Forbes*. [Online]: https://www.forbes.com/sites/louiscolumbus/2016/11/27/roundup-of-internet-of-things-forecasts-and-market-estimates-2016/#6a558beb292d. [Accessed on]: 5/9/2019

Davidson, B. I. & Ellis, D. A. (2019). Social media addiction: technological déjà vu. *BMJ*. 365: I4277, DOI: 10.1136/bmj.I4277

Davidson, B., Joinson, A., & Jones, S. L. (2018). Technologically Enhanced Dating: Augmented Human Relationships, Robots, and Fantasy. In Z. Papacharissi (Ed.), *A Networked Self and Love* (pp. 129–155). New York, USA: Routledge. https://doi.org/10.4324/9781315193472

Dellarocas, C. (2003). The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science*, *49*(10), 1407–1424. https://doi.org/10.1287/mnsc.49.10.1407.17308

Dittmar, J. E. (2011). Information technology and economic change: the impact of the printing press. *The Quarterly Journal of Economics*, *126*(3), 1133–1172.

Dolan, E. W. (2018). Strangers smile less to one another when they have their smartphones, study finds. Retrieved March 12, 2019, from https://www.psypost.org/2018/10/strangers-smile-less-to-one-another-when-they-have-their-smartphones-study-finds-52392

Dufresne, Y., & Davidson, B. I. (2019). Big Data (in press). In *Key Concepts in Research Methods*. Oxford, UK: Oxford University Press.

Earp, B. D., & Tradimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*, 621. https://doi.org/10.3389/fpsyg.2015.00621

Edwards, D., & Potter, J. (1992). *Discursive psychology*. London | Newbury Park | New Delhi: Sage.

Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2019.03.006

Ellis, D. A., & Davidson, B. I. (2019). Digital detoxes are a solution looking for a problem. Retrieved March 12, 2019, from https://theconversation.com/digital-detoxes-are-a-solution-looking-for-a-problem-106460

Ellis, D. A., Davidson, B. I., & Kaye, L. K., (2019). Should smartphones be banned for children or does cyberpsychology have a bigger problem? *Cyberpsychology, Behavior, and Social Networking*. [In Press]

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *Special Issue on Human Accuracy in Mobile Data*

Collection in the International Journal of Human-Computer Studies, 127. DOI: 10.1016/j.ijhcs.2019.05.004

Facebook. (2014). Facebook to Acquire Whatsapp. Retrieved January 11, 2018, from https://newsroom.fb.com/news/2014/02/facebook-to-acquire-whatsapp/

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, *58*(2), 203–210.

Field, H. (2003). Causation in a physical world. In M. J. Loux & D. W. Zimmerman (Eds.), *Oxford handbook of metaphysics* (pp. 435–460). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199284221.001.0001

Fioretti, J. (2017). French Privacy watchdog raps Whatsapp over Facebook data sharing. Retrieved January 11, 2018, from https://uk.reuters.com/article/us-whatsapp-privacy-france/french-privacy-watchdog-raps-whatsapp-over-facebook-data-sharing-idUKKBN1EC285

Fiske, S. T. (2010). *Social Beings* (2nd ed.). NJ, USA: Wiley.

Fox, D., Prilleltensky, I., & Austin, S. (2009). Critical Psychology for Social Justice: Concerns and Dilemmas. In D. Fox, I. Prilleltensky, & S. Austin (Eds.), *Critical Psychology An Introduction* (Second Edi, pp. 3–20). Los Angeles | London | New Delhi | Singapore | Washington DC: Sage.

Gangadharbatla, H. (2008). Facebook me: Collective self-esteem, need to belong, and internet self-efficacy as predictors of the iGeneration's attitudes toward social networking sites. *Journal of Interactive Advertising*, *8*(2), 5–15.

Geels, F. W. (2007). Feelings of Discontent and the Promise of Middle Range Theory for STS: Examples from Technology Dynamics. *Science, Technology, & Human Values, 32(6)*, 627-651

Gilman, A., McC, R., Sestieri, A. M. B., Cazzella, A., Claessen, H. J., Cowgill, G. L., … Harrison, R. J. (1981). The Development of Social Stratification in Bronze Age Europe [and Comments and Reply]. *Current Anthropology*, *22*(1), 1–23.

Herrmann, T., Jahnke, I., & Loser, K. U. (2004). The Role Concept as a Basis for Designing Community Systems. *Cooperative Systems Design, Scenario-Based Design of Collaborative Systems*, 163–178.

Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, *30*(6), 377–386. https://doi.org/10.1177/0270467610385893

Hogg, M., Terry, D., & White, K. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*. https://doi.org/10.2307/2787127

internet live stats. (2019). Twitter Usage Statistics. Retrieved March 4, 2019, from http://www.internetlivestats.com/twitter-statistics/

Joinson, A. N. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers*, *31*(3), 433–438.

Joinson, A. N., (2001). Self-disclosure in computer-medicated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, 31, 177-192

Joinson, A. N., & Paine, C. B., (2007). Self-disclosure, privacy and the internet. *Oxford Handbook of Internet Psychology*, 3, 237-252

Jones, S. L., Ferreira, S., Hosio, S., Goncalves, J., & Kostakos, V., (2015). Revisitation analysis of smartphone app use. *In the proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* Osaka, Japan. 1197-1208

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292.

Kühn, S., Kugler, D. T., Schmalen, K., Weichenberger, M., Witt, C., & Gallinat, J. (2018). Does playing violent video games cause aggression? A longitudinal intervention study. *Molecular Psychiatry*, 1–15. https://doi.org/10.1038/s41380-018-0031-7

Kuijer, L., & Giaccardi, E. (2018). Co-performance: Conceptualizing the Role of Artificial Agency in the Design of Everyday Life. In *In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 125–139).

Kushlev, K., Hunter, J. F., Proulx, J., Pressman, S. D., & Dunn, E. (2019). Smartphones reduce smiles between strangers. *Computers in Human Behavior*, *91*, 12–16. https://doi.org/10.1016/j.chb.2018.09.023

Latour, B., (2007). Beware, your imagination leaves digital traces. *Times Higher Literary Supplement*, 6(4), 129-131

Levina, N., & Arriaga, M. (2014). Distinction and Status Production on User-Generated Content Platforms: Using Bourdieu's Theory of Cultural Production to Understand Social Dynamics in Online Fields. *Information Systems Research*, *25*(3), 468–488.

Levine, M., Cassidy, C., & Brazier, G. (2002). Self-Categorization and Bystander

Non-Intervention: Two Experimental Studies. *Journal of Applied Social Psychology*, *32*(7), 1452–1463.

Levine, M., & Crowther, S. (2008). The responsive bystander: How social group membership and group size can encourage as well as inhibit bystander intervention. *Journal of Personality and Social Psychology*, *95*(6), 1429–1469.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems*, *24*, 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Mackay, H., & Gillespie, G. (1992). Extending the social shaping of technology approach: ideology and appropriation. *Social studies of science*, 22(4), 685-716.

MacKenzie, D., & Wajcman, J. (1999). The Social Shaping of Technology. 2nd Edition. Open University Press. Buckingham, UK. ISBN: 9780335199136

Manogaran, G., Thota, C., & Lopez, D. (2018). Human-computer interaction with big data analytics. In *HCI challenges and privacy preservation in big data security*. GIG Global. 1-12

Marder, B., Houghton, D., Joinson, A., Shankar, A., & Bull, E. (2016). Understanding the psychological process of avoidance-based self-regulation on Facebook. *CyberPsychology, Behavior & Social Networking*, *19*(5), 321–327. https://doi.org/10.1089/cyber.2015.0564

Marder, B., Joinson, A., & Shankar, A. (2012). Every post you make, every pic you take, I'll be watching you: Behind social spheres on facebook. *Proceedings of the Annual Hawaii International Conference on System Sciences*, (July 2015), 859–868. https://doi.org/10.1109/HICSS.2012.12

Marder, B., Joinson, A., Shankar, A., & Houghton, D. J. (2016). The Extended 'Chilling' Effect ofFacebook: The Cold Reality of Ubiquitous Social Networking. *Computers in Human Behavior*, *60*, 582–592. Retrieved from http://opus.bath.ac.uk/49558/1/Chilling_Effect_CHB_revised_submission_BM.pdf

Marder, B., Shankar, A., Houghton, D., & Joinson, A. (2017). 'What if my mum sees it?': Examination of visible brand interaction in the presence of a wider network. *Information Technology & People*, *30*(1), 210–226. https://doi.org/10.1108/ITP-10-2015-0263

Marwick, A. E., & boyd, danah. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media and Society*, *13*(1), 114–133. https://doi.org/10.1177/1461444810365313

McCarney, R., Warner, J., Iliffe, S., Van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC medical research methodology*, *7*(1), 30.

McDonald, T. (2018). iFrown: Having a mobile phone makes us less likely to smile because they make people less engaged with their social environment. Retrieved March 12, 2019, from https://www.dailymail.co.uk/news/article-6223185/iFrown-Having-mobile-phone-stops-smiling-changes-way-behave.html

McKenna, K. Y. A., & Bargh, J. A., (1998). Coming Out in the Age of the Internet: Identity "Demarginalization" Through Virtual Group Participation. *Journal of Personality and Social Psychology*, 75(3), 681-694

Mitzner, T. L., Boron, J. B., Fausset, C. B., Adams, A. E., Charness, N., Czaja, S. J., … Joseph, S. (2010). Older Adults Talk Technology: Technology Usage and Attitudes. *Computers in Human Behavior*, *26*(6), 1710–1721. https://doi.org/10.1016/j.chb.2010.06.020

Oliver, E., (2010). The challenges in large-scale smartphone user studies. *In proceedings of the 2nd ACM International Workshop on Hot Topics in Planet-scale Measurement – HotPlant'10*. https://doi.org/10.1145/1834616.1834623

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776-783

Papacharissi, Z. (2002). Presentation of Self In Virtual Life: Characteristics of Personal Home Pages. *Journalism & Mass Communication Quarterly*, 643–660.

Parsons, T. D. (2015). Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective, and Social Neurosciences. *Frontiers in Human Neuroscience*, https:doi.org/10.3389/fnhum.2015.00660

Pfeil, U., Svangstu, K., Ang, C. S., & Zaphiris, P. (2011). Social Roles in an Online Support Community for Older People. *International Journal of Human-Computer Interaction*, *27*(4), 323–347. https://doi.org/10.1080/10447318.2011.540490

Preece, J., & Schneiderman, B. (2009). The Reader-to-Leader Framework: Motiviating Technology-Mediated Social Participation. *Transactions on Human-Computer Interaction*, *1*(1), 1–25. https://doi.org/10.5121/ijfcst.2014.4403

Przybylski, A. K., & Weinstein, N. (2019). Violent video game engagement is not associated with adolescents' aggressive behavior: evidence from a registered report. *Royal Society Open Science*, *6*(2). https://doi.org/10.1098/rsos.171474

Roser, M., & Ritchie, H., (2019). Technological Progress. *OurWorldInData.org*. [Online]: https://ourworldindata.org/technological-progress. [Accessed on]: 5/9/2019

Rui, J., & Stefanone, M. A. (2013). Strategic self-presentation online : A cross-cultural study. *Computers in Human Behavior*, *29*(1), 110–118. https://doi.org/10.1016/j.chb.2012.07.022

Sawyer, J., (2019). What Social Word Got Right and Why It's Needed for our [Technology] Revolution. *Journal of Design and Science*. [Online]: https://jods.mitpress.mit.edu/pub/sawyer. [Accessed on]: 5/6/2019

Shaw, H., Ellis, D. A., & Ziegler, F. V. (2018). The Technology Integration Model (TIM). Predicting the continued use of technology. *Computers in Human Behavior*, *83*, 204–214. https://doi.org/10.1016/j.chb.2018.02.001

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, *69*, 487–510. https://doi.org/10.1146/annurev-psych-122216- 011845

Stroebe, W., & Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. *Perspectives on Psychological Science*, *9*(1), 59–71. https://doi.org/10.1177/1745691613514450

Syed, S., Voelske, M., Potthast, M., & Stein, B. (2018). Dataset for generating TL;DR [Dataset]. https://doi.org/10.5281/zenodo.1168855

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, *211*(4481), 453–458.

van Dijck, J. (2013). 'You have one identity': performing the self on Facebook and LinkedIn. *Media, Culture & Society*, *35*(2), 199–215. https://doi.org/10.1177/0163443712468605

Vasalou, A., & Joinson, A. N. (2009). Me, myself and I: The role of interactional context on self-presentation through avatars. *Computers in Human Behavior*, *25*(2), 510–520. https://doi.org/10.1016/j.chb.2008.11.007

Wachter, S., & Mittelstadt, B., (2018). A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*. [In Press]

Weaver, S. D., & Gahegan, M. (2007). Constructing, Visualizing and Analyzing a Digital Footprint. *Geographical Review*, *97*(3), 324–350.

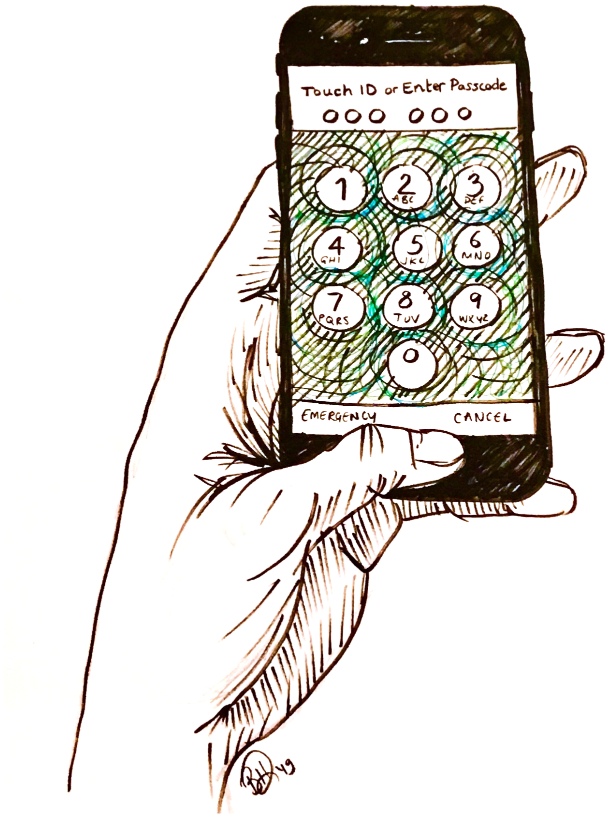
Welser, H., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. *Proceedings of the 2011 IConference on - IConference '11*, 122–129. https://doi.org/10.1145/1940761.1940778

Welser, H. T., Gleave, E., Barash, V., Smith, M., & Meckes, J. (2009). Whither the Experts?: Social affordances and the cultivation of experts in community Q&A systems. *Proceedings - 12th IEEE International Conference on Computational Science and Engineering, CSE 2009*, *4*(October 2016), 450–455. https://doi.org/10.1109/CSE.2009.254

Zhao, S., Ramos, J., Tao, J., Jiang, Z., Wu, Z., …, Dey, A. K., (2016). Discovering different kinds of smartphone users through their application usage behaviors. *In proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany.* 498-509

CHAPTER II

DO SMARTPHONE USAGE SCALES PREDICT BEHAVIOR?



Chapter Publication Status:

Ellis, D. A., **Davidson, B. I.**, Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? Special Issue: Human Accuracy in Mobile Data Collection in the *International Journal of Human-Computer Studies*. 130. pp. 86-92. DOI: 10.1016/j.ijhcs.2019.05.004

Chapter uses a **post-print** (open access) version found here:

https://psyarxiv.com/6fjr7/

Statement of Authorship (to preface each co-authored paper)

<table>
<tr><td colspan="10">This declaration concerns the article entitled:</td></tr>
<tr><td colspan="10">Do smartphone usage scales predict behavior?</td></tr>
<tr><td colspan="10">Publication status (tick one)</td></tr>
<tr><td>Draft manuscript</td><td></td><td>Submitted</td><td></td><td>In review</td><td></td><td>Accepted</td><td></td><td>Published</td><td>X</td></tr>
<tr><td>Publication details (reference)</td><td colspan="9">Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K., (2019), Do smartphone usage scales predict behavior? Special Issue on Human Accuracy for the International Journal of Human-Computer Studies, 130C, DOI: 10.1016/j.ijhcs.2019.05.004

Post-print available online: https://psyarxiv.com/6fjr7/</td></tr>
<tr><td colspan="10">Copyright status (tick appropriate statement)</td></tr>
<tr><td colspan="4">I hold the copyright for this material</td><td></td><td colspan="4">Copyright is retained by the publisher, but I have been given permission to replicate the material here</td><td>X</td></tr>
<tr><td>Candidate's contribution to the paper (detailed, and also given as a percentage).</td><td colspan="9">Formulation of ideas:
DAE, BID, HS, KG [40; 35; 20; 5]

Design of methodology [e.g., survey production]:
DAE, BID, HS [33; 33; 33]

Data collection:
BID, HS, DAE [60; 35; 5]

Data analysis
BID, DAE [90; 10]
R code written by BID with input from DAE

Data Visualization
BID [100]

Presentation of data in journal format:
BID [60], DAE [40] wrote first version of manuscript. All authors contributed to revisions.</td></tr>
<tr><td>Statement from Candidate</td><td colspan="9">This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.</td></tr>
<tr><td>Signed</td><td colspan="6">Brittany Davidson</td><td>Date</td><td colspan="2">27th June 2019</td></tr>
</table>

Chapter I provided an overview of the forthcoming thesis and chapters. The main interest of this thesis relates to behavior measurement – and considering what different methods and approaches can reveal about individuals and groups in terms of their behavior (e.g., usage, communication style) and identity. This chapter starts to address the second research question of this thesis: 'What can we understand about an individual's interactions with technologies?', where we sought to test how reliable and accurate traditional technology usage proxies are.

The following chapter provides a brief overview of 'technology usage' measurement within social psychology. Typically, this is measured via self-report, which is unlikely to be adequate as many technologies (e.g., smartphones, laptops) are highly integrated into daily life and technology usage, particularly smartphone usage, has arguably become second nature (e.g., Boase & Ling, 2013; Doughty et al., 2012; Jungselius & Weilenmann, 2018). Hence, there is a question as to whether individuals are able to accurately report their technology usage via self-reported measures or providing usage estimates (Shaw, Ellis, & Ziegler, 2018), which we examine here.

The present chapter focuses on iPhone usage specifically. We selected and employed ten popular self-report measures used across research to measure 'smartphone use' (e.g., Smartphone Addiction Scale (SAS), Smartphone Application-Based Addiction Scale (SABAS), Problematic Mobile Phone Use Questionnaire (PMPUQ)), alongside asking participants to estimate their smartphone usage, and collected objective usage data form their phones in order to understand whether these smartphone usage scales (or estimates) can predict behavior.

This piece acts as a validation study, which demonstrated that these scales and estimates are inadequate proxies for behavior, and therefore, we should utilize technologies (e.g., Apple Screen Time) to enrich and improve research moving forward if the core research question relates to *actual* behavior. We would expect the same disparity between scales measuring other technology or social media usage, where research should utilize computational methods and digital traces to better understand technology usage (see *Chapter III* and *IV*). However, as noted in *Chapter I*, the research question is critically important, where we need to distinguish between 'experiences' with and 'behavior' measurement with technology (*Chapter V*). This also has a wider impact as the ‘screen time debate’ continues to discuss the impacts of screen time on children and society more generally (UK Parliament, 2018) when much of the prior work has not actually measured ‘behavior’ and instead relied on inadequate proxies. Hence, utilizing new technologies and tools is essential moving forward.

Abstract

Understanding how people use technology remains important, particularly when measuring the impact this might have on individuals and society. However, despite a growing body of resources that can quantify smartphone use, research within psychology and social science overwhelmingly relies on self-reported assessments. These have yet to convincingly demonstrate an ability to predict objective behavior. Here, and for the first time, we compare a variety of smartphone use and ‘addiction’ scales with objective behaviors derived from Apple’s Screen Time application. While correlations between psychometric scales and objective behavior are generally poor, single estimates and measures that attempt to frame technology use as habitual rather than ‘addictive’ correlate more favorably with subsequent behavior. We conclude that existing self-report instruments are unlikely to be sensitive enough to accurately predict basic technology use related behaviors. As a result, conclusions regarding the psychological impact of technology are unreliable when relying solely on these measures to quantify typical usage.

1 Introduction

1.1 Background

Despite decades of progress, understanding the overall impact of technology on people and society remains a challenge (Shaw et al., 2018). Perhaps this is because such a topic naturally aligns itself with many disparate research questions. Investigations range from issues concerning problematic use (e.g., can smartphones disrupt sleep?), to the effects of engaging with feedback as part of a behavior change intervention (e.g., does monitoring physical activity improve health?) (Ellis & Piwek, 2018). Approaches to date in behavioral science have almost exclusively focused on asking people to consider their personal experience with a technology in order to better understand its impact (Ellis, Kaye, Wilcockson, & Ryding, 2018). This mirrors a general trend within social psychology as a whole (Baumeister, Vohs, & Funder, 2007; Dolinski, 2018), but it is perhaps more surprising when applied to mobile and pervasive systems that can record human-computer interactions directly (Piwek, Ellis, & Andrews, 2016). Smartphones have provided several new opportunities in this regard (Miller, 2012). For example, behavioral interactions can be measured 'in situ' with a variety of applications and those in computer science have been measuring these interactions for several years (Jones, Ferreira, Hosio, Goncalves, & Kostakos, 2015; Oliver, 2010; Zhao et al., 2016). However, methodological developments have had very little impact on how the majority of social science attempts to quantify, explain, and understand technology use more generally.

Two common methods are often deployed by social scientists to capture technology usage 'behaviors'. The first relies on participants providing estimates of frequency or duration (Butt & Phillips, 2008). However, this method has previously been described as 'sub-optimal' when attempts are made to validate single measures against objective behavior (e.g., Boase & Ling, 2013). In addition, the use of multiple technologies simultaneously (e.g., a smartphone and a laptop) mean that these estimates have become even more problematic due the level of cognitive burden required to quantify many different types of habitual behavior (Boase & Ling, 2013; Doughty, Rowland, & Lawson, 2012; Jungselius & Weilenmann, 2018). In response to these critiques, a second method utilizes questionnaires that aim to quantify technology related experiences. Considering smartphones specifically, an abundance of self-reported measures have been created in an attempt to capture and predict actual behavior (e.g., Bianchi & Phillips, 2005; Billieux, Van Der Linden, & Rochat, 2008; Csibi, Demetrovics, & Szabó, 2016; Kwon, Kim, Cho, & Yang, 2013; Rosen, Whaling, Carrier, Cheever, & Rokkum, 2013; Sivadas & Venkatesh, 1995; Yildirim & Correia, 2015). Following traditional methods associated with scale development, factor analyses ensure that such assessments are reliable, but less emphasis has been placed on establishing validity. This sets these scales apart from other areas where self-report has been rigorously validated against behavioral metrics (e.g., personality) (e.g., McCrae & Costa, 1987; Parker & Stumpf, 1998). The lack of validation and clarity regarding constructs and measurement is therefore detrimental to the sound utilization of these scales in subsequent research (Clark & Watson, 1995).

Many measures are conceptualized around 'smartphone behaviors', and are used by many researchers to provide a proxy measure of usage (Ellis et al., 2018). Perhaps more importantly, research utilizing these assessments tends to use high-scores to correlate smartphone usage with a variety of negative outcomes (e.g., depression and anxiety) (e.g., Elhai, Dvorak, Levine, & Hall, 2017; Richardson, Hussain, & Griffiths, 2018) and provide evidence for the classification of a behavioral addiction (e.g., Tao et al., 2017; Wolniewicz, Tiamiyu, Weeks, & Elhai, 2018). This repeats a pattern of research priorities that previously focused on the negative impacts of many other screen-based technologies, systematically moving from television and video games, to the internet and social media (Przybylski & Weinstein, 2017; Rosen et al., 2014). However, the few studies that have measured behavior directly, tend to demonstrate conflicting results. For example, Rozgonjuk et al. (2018) observed no association between smartphone use and severity of depression or anxiety. Further, higher levels of reported depression correlated with individual's checking their phone *less* over a week. Therefore, the notion of reducing 'screen time' and technology may be counter-intuitive, as a sudden reduction in smartphone use may in fact be an early warning sign of social withdrawal (Mou, 2016).

1.2 The Present Study

To date, only a handful of small studies have attempted to validate these scales in small samples that focus on single measures with mixed results (Andrews, Ellis, Shaw, & Piwek, 2015; Elhai et al., 2018; Foerster, Roser, Schoeni, & Röösli, 2015; Lin, Chiang, & Jiang, 2015; Rozgonjuk et al., 2018; Wilcockson, Ellis, & Shaw, 2018). Here, we attempt to compare the human accuracy of ten smartphone usage scales and single estimates against objective measures of smartphone behavior. This takes advantage of a recent iOS update from Apple, which automatically logs a series of behavioral metrics related to 'screen time' over a period of seven days. Data available includes the length of time users spend on their devices, the number of times the phone is picked up, alongside the number of notifications received daily. This allowed for several attempts at validation that includes correlations and cluster-based analyses. The latter of which compares the overlap between high-usage groups derived independently from self-report scores or behavioral metrics.

2 Method

2.1 Ethics

This study was ethically approved by the University of Bath School of Management (ID: 2392) and was conducted in accordance with guidelines provided by the British Psychological Association (BPS).

2.2 Participants

Participants were recruited from within affiliated universities (Lancaster, Bath, and Lincoln) (23.12%), or using the Prolific Academic platform (76.89%). Participants were paid a small sum for their participation via Prolific Academic (£5.34/hr) and provided informed consent. 238 participants (124 female, mean age = 31.88; *SD* = 11.19) who owned an iPhone 5 or above and had been running the latest version of iOS for at least one week were eligible to participate. Our sample size is comparatively larger than other studies that have previously attempted to validate these scales and includes data from a comparable time frame (Andrews et al., 2015; Elhai et al., 2018; Lin et al., 2015; Rozgonjuk et al., 2018; Wilcockson et al., 2018). In addition, our sample is similar to studies that utilize these scales when making links between smartphone use and other correlates, for example, Wolniewicz et al (2018), N=296 and Elhai, Levine, Dvorak, and Hall (2016), N = 308.

2.3 Procedure and Materials

All participants were directed to a Qualtrics survey hosted by the University of Lincoln. Participants first provided an estimate of how many hours and minutes they spend on their iPhone daily. They were also asked to estimate the number of notifications received daily, and how many times they pick up their device each day.

Next, they completed ten scales that aim to asses smartphone usage and/or associated constructs (Table 1). Scales were selected based on their popularity and broad range of conceptualizations (e.g., attachment, fears, 'addictions', etc.) and were presented at random within the survey. Finally, participants transferred their latest Screen Time capture data from Apple's Screen Time app to provide the actual number of hours and minutes spent on their phone, number of notifications received, and number of times they had picked up their device each day for a period of one week. Daily averages were calculated for all three behavioral metrics.

Mobile Phone Problem Use Scale (MPPUS)
(Bianchi & Phillips, 2005)
The MPPUS is a 27-item scale designed to assess problematic usage of mobile phones, with each item scored via a Likert scale ranging from 'Not true at all' (1) to 'Extremely true' (10). Higher scores denote increased levels of problematic usage.

Nomophobia Questionnaire (NMP-Q)
(Yildirim & Correia, 2015)
The NMP-Q is a 20-item designed to assess nomophobia. This is defined as a phobia of being separated from one's smartphone. Each statement is scored using a 7-point Likert scale from 'Strongly disagree' (1) to 'Strongly agree' (7). Higher scores correspond to higher nomophobia severity, where scores of <20 denote an absence of nomophobia, >20 – <60 denotes mild nomophobia, >=60 – <100 denotes moderate nomophobia, with scores >= 100 suggesting severe nomophobia.

Possession Incorporation in the Extended Self
(Sivadas & Venkatesh, 1995)
This scale comprises of 6-items that aims to determine the extent possessions have become incorporate into an 'extended self' originally defined by Belk (1988). Statements are scored using a 7-point Likert scale ranging from 'Strongly disagree' (1) to 'Strongly agree' (7). We used the specific-possession incorporation version, where the items were phrased as follows: 'x helps me achieve the identity I want to have', with x substituted as 'my smartphone,'. Higher scores denote an increased integration of a smartphone an identity.

Attachment Scale

(Sivadas & Venkatesh, 1995)

The attachment scale contains 4-items, which aims to assess the attachment to an object, in this case a smartphone, for example, 'I am emotionally attached to my smartphone'. This used a 7-point Likert scale ranging from 'Strongly disagree' (1) to 'Strongly agree' (7). Higher scores correspond to higher levels of attachment to the object in question.

Smartphone Addiction Scale (SAS)

(Kwon et al., 2013)

The SAS is a 33-item scale designed to measure smartphone 'addiction', with each statement scored via a 6-point Likert scale from 'Strongly disagree' (1) to 'Strongly agree' (6). It consists of six factors: daily life disturbance, positive anticipation, withdrawal, cyberspace-orientated relationship, overuse, and tolerance. These can be combined to provide a single score. Higher scores correspond to higher smartphone usage and 'addiction'.

Smartphone Application-Based Addiction Scale (SABAS)

(Csibi et al., 2016)

We used the English version of the SABAS scale, which comprises of 6-items, with each item scored using 6-point Likert scale from 'Strongly disagree' (1) to 'Strongly agree' (6). It aims to assess application-based addictions associated with smartphones. Higher scores correspond to higher smartphone (application) usage and 'addiction'.

Problematic Mobile Phone Use Questionnaire (PMPUQ)

(Billieux et al., 2008)

The PMPUQ aims to assess actual and potential problematic usage of mobile phones. We used a short 15-item version, which concerned mobile phone usage when driving, forbidden use of mobile phones, and use of mobile phones in dangerous situations. The scale is traditionally a 4-item Likert scale from 'Strongly disagree' (1) to 'Strongly agree' (4), however, we also included an additional 'Not Applicable' (5) for those who did not drive in our sample (coded as 0). Higher scores correspond with increased levels of problematic usage.

Media and Technology Usage and Attitudes Scale (MTUAS)

(Rosen et al., 2013)

The complete MTUAS comprises of 66-items that aims to assess technology and media use more widely. However, here we used 9-items from a subscale, which focuses on smartphone use (items 9-17). Each item is scored on a 10-point scale from 'Never' (1) to 'All the time' (10), where the mean measure is taken for each participant. Higher means correspond to higher smartphone usage.

Smartphone Use Questionnaires (SUQ-G&A)

(Marty-Dugas, Ralph, Oakman & Smilek, 2018)

SUQ-G&A seeks to distinguish general smartphone usage and absent-minded smartphone usage. This provides scores from two 10-item scales: general (SUQ-G) and absent-minded (SUQ-A). Both use a 7-point scale from 'Never' (1) to 'All the time' (7). SUQ-G focusses on specific uses, e.g., 'How often do you check social media apps such as Snapchat, Facebook, or Twitter', and the SUQ-A asks questions regarding mindless usage, e.g., 'How often do you find yourself checking your phone without realizing why you did it?'. Higher mean scores correspond to higher smartphone usages (general or absent-minded).

2.4 Analysis Plan

Scores for each scale were calculated (as detailed above), with manipulations for reversed items as necessary. Tables 1 and 2 provide descriptive statistics for all self-reported and behavioral metrics. Pearson's Correlations (Table 3) were calculated between all self-reported measures, single estimates, and objective behavioral metrics. While we note that the average number of notifications is not strictly a behavioral measure, it is included here to provide context regarding how often a person may be expected to pick up or check their phone as notifications act as a request for user attention. Therefore, this provides an additional validity check as we expect to observe a positive correlation between the number of notifications and the amount of time a person spends on their phone. The overall performance of each self-report measure was derived from the mean correlation across all three objective behavioral measures (Figure 1). For example, the mean score for a single duration estimate was based on mean of three correlations between the estimate and behavioral averages of (1) hours use, (2) pickups, and (3) notifications. Finally, a series of k-means algorithms considered overlaps in classification when participants were clustered using only self-report or objective behavior (Figure 2).

3 Results

3.1 Self-Reported Measures

Table 1 reports the means, standard deviations, and internal consistency measures (Cronbach's Alpha (α) for all self-reported measures.

Table 1. Descriptive Statistics (means (M) and standard deviations (SD)) for single estimates and self-report assessments. Highest and lowest possible scores for each measure are provided for reference.

Self-report measures	Items	Min-max	*M*	*SD*	α
Single time estimate (minutes) (TEst)	1	-	226.6	128.37	
Single pickup estimate (PEst)	1	-	45.69	42.16	
Single notification estimate (NEst)	1	-	39.09	42.46	
Mobile phone problem use scale (MPPUS)	27	27–270	111.90	43.12	.94
Nomophobia scale (NS)	11	20–140	82.57	25.76	.96
Possession incorporation in the extended self (ES)	6	6–42	21.53	8.99	.93
Smartphone attachment scale (SAt)	4	4–24	17.02	6.05	.87
Smartphone addiction scale (SAS)	33	33–198	94.20	30.17	.95
Smartphone application-based addiction scale (SABAS)	6	6–36	15.83	5.89	.81
Problematic mobile phone use questionnaire (PMPUQ)	15	15–60	27.54	5.85	.72
Media and technology usage and attitudes scale (MTUAS)	9	9–90	6.24	1.33	.84
Smartphone use questionnaire (general) (SUQ-G)	10	10–70	48.45	8.89	.78
Smartphone use questionnaire (absent minded) (SUQ-A)	10	10–70	45.60	14.37	.95

3.2 Behavioral Metrics

Table 2 presents means and standard deviations from objective behavioral measures. Data were available for the previous seven days, however, the day of data collection is naturally incomplete, so all behavioral metrics are based on an average from six complete days of data from each participant. Previous research has suggested that identical smartphone usage collected for a minimum of five days will reflect typical weekly usage, with habitual checking behaviors (pickups) requiring a minimum of two complete days of collection irrespective of weekday (Wilcockson et al., 2018). A series of one-way ANOVAs confirm that no weekday differences were present in any of our behavioral data (all p's > .2). Finally, we note that participants, on average, pickup their phones fewer times when compared to the number of notifications received (1:1.05 ratio of pick ups to notifications).

Table 2. Descriptive Statistics for Behavioral Measures (means (M) and standard deviations (SD)). These are in line with previous research considering smartphone behaviors in smaller samples (e.g., Andrews et al., 2015).

Behavioral Measure	*M*	*SD*
Time (minutes)	232.66	119.44
Pick ups	85.84	53.34
Notifications	90.13	88.86

3.3 Correlations

Pearson's correlation coefficients were calculated across single estimates, self-reported scales, and behavioral data (Table 3). All self-reported scales positively correlated with objective time spent on a smartphone (ObjT). These varied from .40 to .13. However, a single estimate of time (TEst) was a better predictor than any self-report scale [r =.48].

Average number of objective pickups (ObjP) modestly correlated with the Smartphone Usage Questionnaire - General (SUQ-G) [r = .31] and Smartphone Usage Questionnaire – Absent Minded (SUQ-A) [r = .30]. Weak correlations were observed between the Smartphone Addiction Scale (SAS) [r = .22], Mobile Phone Problem Use Scale (MPPUS) [r = .18], and Media and Technology Usage and Attitudes Scale (MTUAS) [r = .15]. Again, a single estimate of pickups (PEst) was a superior predictor in comparison to any self-report instrument [r = .32].

Average number of notifications (ObjN) weakly correlated with most self-reported scales (exceptions are the Extended Self (ES), Smartphone Application Application-Based Addiction Scale (SABAS), and the Problematic Mobile Phone Use Questionnaire (PMPUQ)). These varied from .28 to .15. A single estimate of daily notifications received (NEst) correlated moderately with the objective counterpart (ObjN) [r = .53].

Table 3. Pearson's correlations between single estimates, self-reported scales, and objective behavior.

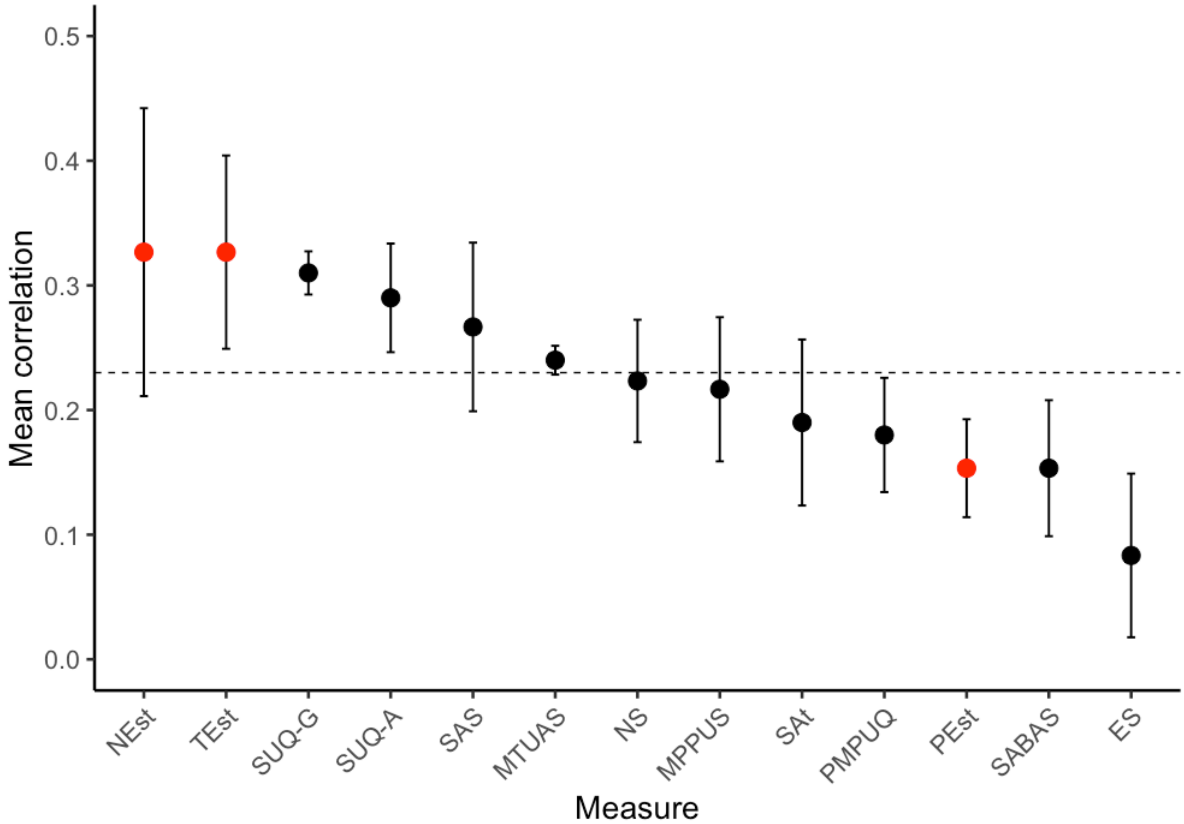
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Age																
2. TEst	-.22**															
3. PEst	-.10	.22**														
4. NEst	-.15*	.30**	.32**													
5. MPPUS	-.08	.28**	.14*	.06												
6. NS	-.03	.22**	.08	.06	.74**											
7. ES	.14*	.14*	.07	.00	.53**	.56**										
8. SAt	.02	.21**	.04	.03	.46**	.54**	.69**									
9. SAS	-.08	.29**	.09	.06	.82**	.75**	.62**	.59**								
10. SABAS	-.03	.21**	.13	.05	.77**	.68**	.55**	.52**	.76**							
11. PMPUQ	-.04	.27**	.17**	.14*	.55**	.46**	.38**	.37**	.56**	.48**						
12. MTUAS	-.26**	.28**	.24**	.22**	.36**	.38**	.23**	.32**	.34**	.25**	.37**					
13. SUQ-G	-.28**	.36**	.14*	.24**	.56**	.54**	.39**	.41**	.57**	.43**	.42**	.60**				
14. SUQ-A	-.26**	.24**	.14*	.04	.66**	.58**	.35**	.40**	.62**	.53**	.47**	.45**	.69**			
15.ObjT	-.20**	.48**	.10	.13*	.33**	.32**	.21**	.32**	.40**	.26**	.27**	.26**	.34**	.36**		
16. ObjP	-.32**	.23**	.23**	.32**	.18**	.16*	-.01	.10	.22**	.12	.15*	.24**	.31**	.30**	.39**	
17. ObjN	-.35**	.27**	.13*	.53**	.14*	.19**	.05	.15*	.18**	.08	.12	.22**	.28**	.21**	.37**	.66**

Note: *Correlation is significant at a .05 level (2-tailed) **Correlation is significant at a .01 level (2-tailed)

TEst = Single time estimate, PEst = Single pick-up estimate, NEst = Single notification estimate, MPPUS = Mobile phone problematic use scale, NS = Nomophobia scale, ES = Possession incorporation in the extended self, SAt = Smartphone attachment, SAS = Smartphone addiction scale, SABAS = Smartphone application-based addiction scale PMPUQ = Problematic mobile phone use questionnaire, MTUAS = Media and technology usage and attitudes scale, SUQ-G = Smartphone use questionnaire (general), SUQ-A = Smartphone use questionnaire (absent minded), ObjT = Objective average daily screen-time, ObjP = Objective average daily number of pickups, ObjN = Objective average daily number of notifications.

In order to assess which estimates or measures performed the best when predicting behavior in general, we calculated the average correlation from all three objective measures (average time spent on their smartphone, average number of pickups, and average number of notifications), for each self-reported measure, and the three single estimates. From this, we note that the notification (NEst) [r = .33] and time (TEst) [r = .33] estimates had the highest average correlation with the three objective behavioral measures, closely followed by the Smartphone Usage Questionnaire – General (SUQ-G) [r = .31] and Smartphone Usage Questionnaire – Absent Minded (SUG-A) scales [r = .29] (Figure 1).

Figure 1. Average r value for each subjective measure across all three objective behavioral measures. Error bars illustrate standard error. Red indicates a single behavioral estimate. Dotted line represents mean correlation across all measures. Refer to Table 1 for abbreviations.

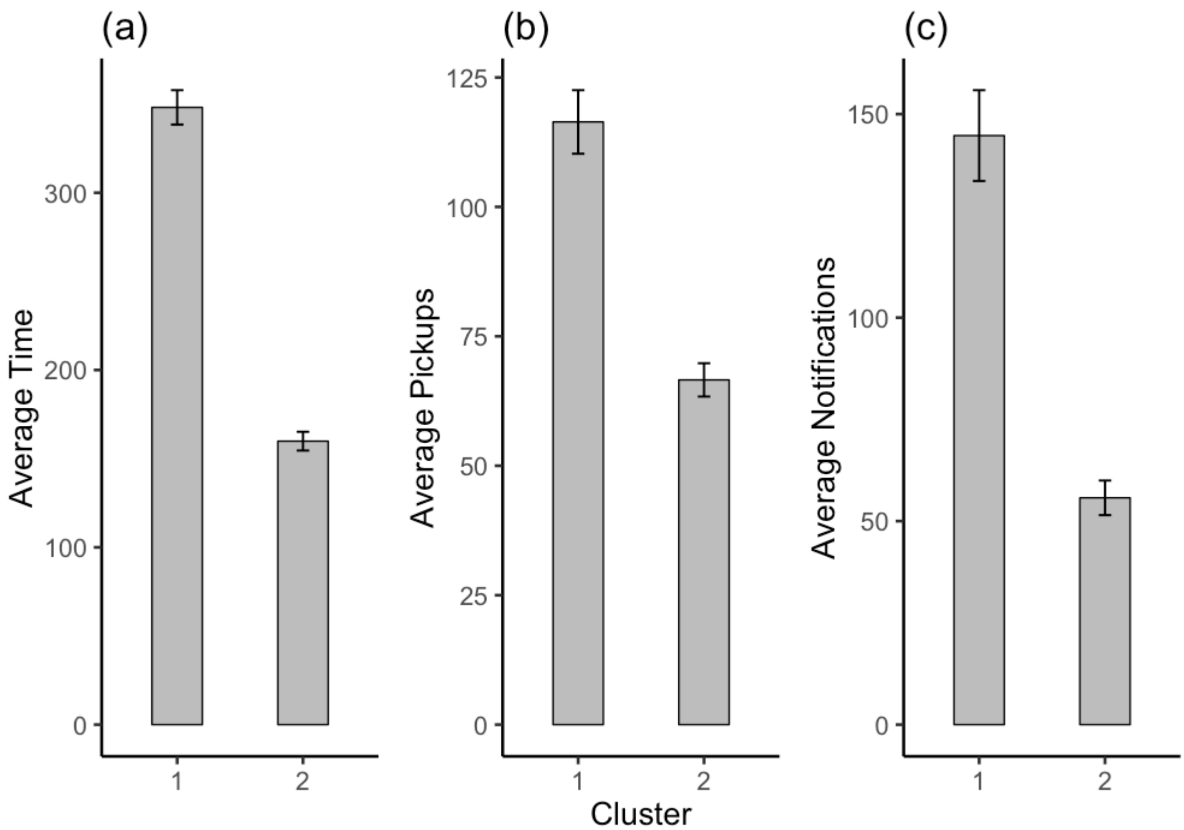


3.4 Cluster Analysis

Many conceptualizations of smartphone use focus on a binary classification whereby 'addiction' or 'problematic' usage are either present or absent. This is also important from a clinical standpoint as these scales are often referred to as having a (potential) diagnostic ability (Lin et al., 2016). Therefore, our final analysis considered if behavioral and self-report measures could classify identical participants. While several unsupervised methods can cluster participants, k-means is widely used in behavioral analytics (e.g., Arazy et al., 2017; Jackson, Østerlund, Maidel, Crowston, & Mugar, 2016; Wang, Brede, Ianni, & Mentzakis, 2018) because it can handle a variety of dataset sizes and produce straightforward outputs (Wu et al., 2008). The unsupervised nature of such an approach also removes any researcher bias.

Participants were clustered into two groups (high and low) twice with different input variables used for each classification. The first cluster analysis used *only* the three objective behavioral measures (time spent, notifications, and pickups). As expected, fewer participants scored highly in all three objective behavioral measurements. Figure 2 illustrates the means of high and low clusters for the objective behavioral measures.

Figure 2. Means of high (N = 92) (cluster 1) and low users (N = 146) (cluster 2) derived from objective data following a k-means cluster analysis. Error bars denote standard error.



A second cluster analysis used only self-reported scales (excluding single estimates) to make a similar distinction. Classifications for each participant were then compared. A large level of agreement between self-report and behavior would lead to identical participants being classified as high in both analyses. However, when comparing classifications between the two data-sets, only 52 of 92 (56.52%) participants identified as high users based on behavior, were also classified as high-users from self-report data.

As expected, the behavioral cluster analysis identified a large percentage (38.66%) of our sample as 'high' users. However, this may lack any meaningful specificity given that comparatively few participants are likely to demonstrate exceptionally high usage patterns (Wilcockson et al., 2018). As a result, research relying on self-report alone has considered non-binary approaches by adopting a three-cluster approach (Lepp, Li, Barkley, & Salehi-Esfahani, 2015). We therefore replicated our previous procedure with a three-cluster solution (k = 3), which separated users into low, medium, and high usage groups. Again, we compared clustering decisions derived from self-report and objective behavior. In this instance, the overlap of high users appearing in both clusters fell to 32.36% (10 out of 31). Here, we observe that moving away from a binary classification does not improve performance.

4 Discussion

To date, no systematic approach has attempted to behaviorally validate the growing number of psychometric instruments, which aim to capture technology related behaviors and experiences. Here, we demonstrate that smartphone related assessments are no better than single duration estimates when predicting subsequent behavior. However, as observed elsewhere, even single-item measurements fail to explain much of the variance associated with comparable behaviors (Boase & Ling, 2013). This has wide-ranging consequences for the vast number of studies that rely on these self-reported measures as a proxy measure of behavior.

Every psychometric scale correlated with at least one objective measure, but the strength of these relationships is far from convincing. Existing smartphone 'addiction' scales, for example, correlated poorly with the 'rapid checking' behaviors that one would associate with a behavioral addiction (Andrews et al., 2015;

Rozgonjuk et al., 2018). As these scales struggle to capture simple behaviors, it remains questionable as to how they could effectively measure habitual, atypical, and more complex behavioral patterns. Further, combining multiple scales did not assist in the identification of participants with high usage patterns derived from behavior alone (see Appendix 1 for more detail and analysis). As a consequence, our results have implications for studies that attempt to understand the impacts of smartphones and other screen-based technologies on health and wellbeing. These issues extend to research that has attempted to link a variety of individual differences (e.g., personality) with technology use (e.g., Butt & Phillips, 2008; Horwood & Anglim, 2018; Takao, Takahashi, & Kitamura, 2009). Errors of measurement here are so large that small effects detected in large-scale research involving estimates may be a component of statistical noise or a weak proxy for other psychological constructs (Ellis, 2019).

While the scales under investigation were developed in an effort to capture specific constructs (e.g., addiction or nomophobia), they are frequently used to quantify usage in the general population. This appears to be in direct conflict with a conceptual framework that problematizes usage without considering how typical these behaviors are within the general population. However, recent conceptualizations of usage perhaps hold some promise. The Smartphone Usage Questionnaires (SUQ) (Marty-Dugas & Ralph, 2018), provided the strongest correlations across the board. These consider everyday smartphone use in the context of attentional lapses and mind wandering instead of conceptualizing everyday behavior as 'addictive or 'problematic', which demonstrates the strength in focusing on cognition directly (e.g., attention to and distraction via technology) rather than addiction. These findings also align with recent theoretical models, which argue that technology use over time becomes habitual and more 'absent-minded' (Shaw et al., 2018). Indeed, a growing body of evidence now supports the notion that psychology should start to move away from a behavioral addictions framework when studying technology use (Panova & Carbonell, 2018).

Broadly speaking, technology usage assessments, which vary from television, to internet, online gaming, and more recently, smartphones, rely on extraordinarily similar scales or estimates – substituting device for device as required (Rosen et al., 2014). This similarity problem can also be considered within smartphone usage scales specifically. Despite being developed years apart and around different frameworks or conceptualizations of use (e.g., fear, attachment, or problematic use, etc.), they appear to, in many cases, measure almost identical constructs. The majority of smartphone usage scales by their very nature likely overlap with higher levels of anxiety and depression rather than smartphone usage, as the item's wording tends to be conceptually similar to that of depression and anxiety scales. One future study may wish to compare how these measures correlate with anxiety assessments and objective behavior. Our results suggest that the correlation would be far stronger with the former than the latter.

Given the complexities associated with studying the impact of technology on people and society, there is an urgent need for basic research to consider what this means for different individuals, devices, contexts, and in the case of smartphones, specific types of app usage (Jungselius & Weilenmann, 2018). The discipline may need to consider a paradigm shift, which would also help drive theoretical development and encourage a systematic shift away from the repetitive development of self-report assessments (Billieux, Maurage, Lopez-Fernandez, Kuss, & Griffiths, 2015). However, this may already be changing as Apple and Google are providing more of this data directly to all users, which provides a simple way to capture basic measures of objective behavior. We anticipate that this alone will lead to many other researchers making use of data derived from these screen time applications in the future. All this is not to suggest that there is no place for self-report or psychometric assessment in this domain of research at all. However, psychometric tools should be built around a concrete understanding of what (a) such measures can accurately assess and (b) what specific questions they can answer. For example, while functions of addiction can go beyond use (e.g., craving), the consumption of technology continues to be frequently referenced as a key metric by researchers in this domain (Dowling & Quirk, 2009). There are also certainly more specific behaviors, which might better map onto these psychometric scales, but research to date typically focuses on time spent on a device overall rather than specific sub-sets of behavior (Ellis et al., 2018). This has further implications for smartphone 'addiction' if it were to ever be included as part of the World Health Organization's ICD-11 (2018) alongside gaming disorder, as any diagnostic criteria will almost certainly have to focus on objective behavior, as well as thoughts, attitudes and feelings towards a technology (Lin et al., 2016).

4.1 Limitations

There are some limitations to note. First, while the behavioral measures utilized here are limited (e.g., this study uses daily tracking rather than finer grain temporal measurements based on hourly patterns of usage), we would argue that actually exploring interactions with technology directly provides a more suitable pathway moving forward. A second limitation concerns our specific use of Apple's Screen Time because this system allows participants to view their own data in real-time, which may partly explain why self-reported estimates correlated more favorably with objective behavioral measures. For example, self-reported pickups have previously not shown a relationship with objective behavior in a smaller sample (Andrews et al. (2015). However, the consistency of our results coupled with reminding participants to not look at their devices when providing estimates suggests that an alternative explanation is unlikely. A related issue may concern the omission of Android users, and previous research has suggested that behaviors and personalities differ between iPhone and Android platforms (Shaw, Ellis, Kendrick, Ziegler, & Wiseman, 2016). However, Andrews et al. (2015) reported an almost identical number of daily smartphone pickups (84.68) with a small number of Android users, demonstrating that regardless of operating systems, the average number of pickups reported in our sample remain remarkably similar. Perhaps more importantly, our findings echo earlier validation concerns albeit on a larger scale (Andrews et al., 2015; Elhai et al., 2018; Lin et al., 2015; Rozgonjuk et al., 2018; Wilcockson et al., 2018).

5 Conclusions

Here we attempted to validate smartphone usage scales against a handful of behavioral metrics. Our results suggest that the majority of these self-report smartphone assessments perform poorly when attempting to predict objective smartphone behaviors. Researchers should therefore be cautious when using these measures to link technology use with outcomes concerning health and psychological well-being. They also provide weak evidence to support the development of any diagnostic criteria (e.g., Lin et al., 2016; Tran, 2016). The issues highlighted here feed into a growing consensus that while psychology has acknowledged a problem with replication, the discipline also needs to address similar issues within measurement (Flake & Fried, 2019). Across psychological science, many self-reports remain insufficient for researchers who continue to make large claims, particularly those which pertain to the impact of technology on public health (Boyd & Pennebaker, 2017; Twenge, Joiner, Rogers, & Martin, 2017). We would encourage other researchers where possible, to complement these with objective measures of behavior in order to better understand the impact of technology on people and society more generally.

References

Abdi, H. & Williams, L. J. (2010). Principle component analysis. *WIREs Computational Statistics*, 2, 433-459

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS ONE*, *10*(10), 1–9. https://doi.org/10.1371/journal.pone.0139004

Arazy, O., Lifshitz-Assaf, H., Nov, O., Daxenberger, J., Balestra, M., & Coye, C. (2017). On the "How" and "Why" of Emergent Role Behaviors in Wikipedia. In *In the proceedings of the Conference on Computer-Supported Cooperative Work Conference (CSCW)* (pp. 2039–2051).

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*. https://doi.org/10.1111/j.1745-6916.2007.00051.x

Belk, R. (1988). Possessions and the Extended Self. *Journal of Consumer Research*, *15*(2), 139–168. https://doi.org/10.1086/209154

Bianchi, A., & Phillips, J. G. (2005). Psychological Predictors of Problem Mobile Phone Use. *CyberPsychology & Behavior*. https://doi.org/10.1089/cpb.2005.8.39

Billieux, J., Maurage, P., Lopez-Fernandez, O., Kuss, D. J., & Griffiths, M. D. (2015). Can Disordered Mobile Phone Use Be Considered a Behavioral Addiction? An Update on Current Evidence and a Comprehensive Model for Future Research. *Current Addiction Reports*. https://doi.org/10.1007/s40429-015-0054-y

Billieux, J., Van Der Linden, M., & Rochat, L. (2008). The role of impulsivity in actual and problematic use of the mobile phone. *Applied Cognitive Psychology*. https://doi.org/10.1002/acp.1429

Boase, J., & Ling, R. (2013). Measuring Mobile Phone Use: Self-Report Versus Log Data. *Journal of Computer-Mediated Communication*. https://doi.org/10.1111/jcc4.12021

Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: a new approach to personality in a digital world. *Current Opinion in Behavioral Sciences*, *18*, 63–68. https://doi.org/10.1016/j.cobeha.2017.07.017

Brems, M., (2017). A One-Stop Shop for Principal Component Analysis. *Towards Data Science*. [Online]: https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c. [Accessed on]: 5/9/2019

Butt, S., & Phillips, J. G. (2008). Personality and self reported mobile phone use. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2007.01.019

Clark, L. A., & Watson, D. (1995). Constructing Validity: Basic Issues in Objective Scale Development. *Psychological Assessment*, *7*(3), 309–319.

Csibi, S., Demetrovics, Z., & Szabó, A. (2016). A Rövid Okostelefon Addikció Kérdoív (ROTAK) kidolgozása és pszichometriai validálása iskoláskorú gyermekekkel. TT - [Development and psychometric validation of the Brief Smartphone Addiction Scale (BSAS) with schoolchidren]. *Psychiatr Hung*.

Dolinski, D. (2018). Is psychology still a science of behaviour? *Social Psychological Bulletin*, *13*(2), e25025.

Doughty, M., Rowland, D., & Lawson, S. (2012). Who is on your sofa? TV Audience Communities and Second Screening Social Networks. In *10th European conference on Interactive tv and video*. https://doi.org/10.1145/2325616.2325635

Dowling, N. A., & Quirk, K. L. (2009). No Title. *CyberPsychology & Behavior*, *12*(1), 21–27. https://doi.org/10.1089/cpb.2008.0162

Elhai, J. D., Dvorak, R. D., Levine, J. C., & Hall, B. J. (2017). Problematic smartphone use: A conceptual overview and systematic review of relations with anxiety and depression psychopathology. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2016.08.030

Elhai, J. D., Levine, J. C., Dvorak, R. D., & Hall, B. J. (2016). Fear of missing out, need for touch, anxiety and depression are related to problematic smartphone use. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2016.05.079

Elhai, J. D., Tiamiyu, M. F., Weeks, J. W., Levine, J. C., Picard, K. J., & Hall, B. J. (2018). Depression and emotion regulation predict objective smartphone use measured over one week. *Personality and Individual Differences*, *133*(21–28). https://doi.org/10.1016/j.paid.2017.04.051

Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior*, *97*, 60–66. https://doi.org/10.1016/j.chb.2019.03.006

Ellis, D. A., Kaye, L. K., Wilcockson, T. D. W., & Ryding, F. C. (2018). Digital Traces of Behaviour Within Addiction: Response to Griffiths (2017). *International Journal of Mental Health and Addiction*. https://doi.org/10.1007/s11469-017-9855-7

Ellis, D. A., & Piwek, L. (2018). Failing to encourage physical activity with wearable technology: what next? *Journal of the Royal Society of Medicine*. https://doi.org/10.1177/0141076818788856

Flake, J. K., & Fried, E. I. (2019). Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. *PsyArXiv*. https://doi.org/10.31234/osf.io/hs7wm

Foerster, M., Roser, K., Schoeni, A., & Röösli, M. (2015). Problematic mobile phone use in adolescents: derivation of a short scale MPPUS-10. *International Journal of Public Health*, *60*(2), 277–286. https://doi.org/10.1007/s00038-015-0660-4

Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3), 453-467

Horwood, S., & Anglim, J. (2018). Personality and problematic smartphone use: A facet-level analysis using the Five Factor Model and HEXACO frameworks. *Computers in Human Behavior*, *85*, 349–359. https://doi.org/10.1016/j.chb.2018.04.013

Jaadi, Z., (2019). A step by step explanation of Principal Component Analysis. *Towards Data Science*. [Online]: https://towardsdatascience.com/a-step-by-step-explanation-of-principal-component-analysis-b836fb9c97e2. [Accessed on]: 5/9/2019

Jackson, C., Østerlund, C., Maidel, V., Crowston, K., & Mugar, G. (2016). Which Way Did They Go? Newcomer Movement through the Zooniverse. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, 623–634. https://doi.org/10.1145/2818048.2835197

Jones, S. L., Ferreira, D., Hosio, S., Goncalves, J., & Kostakos, V. (2015). Revisitation analysis of smartphone app use. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1197–1208). Osaka, Japan. https://doi.org/10.1145/2750858.2807542

Jungselius, B., & Weilenmann, A. (2018). Conceptualizing "Use" in Social Media Studies. In *Proceedings of the 9th International Conference on Social Media and Society - SMSociety '18*. https://doi.org/10.1145/3217804.3217938

Kwon, M., Kim, D. J., Cho, H., & Yang, S. (2013). The smartphone addiction scale: Development and validation of a short version for adolescents. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0083558

Lepp, A., Li, J., Barkley, J. E., & Salehi-Esfahani, S. (2015). Exploring the relationships between college students' cell phone use, personality and leisure.

Computers in Human Behavior. https://doi.org/10.1016/j.chb.2014.11.006

Lin, T. T. C., Chiang, Y.-H., & Jiang, Q. (2015). Sociable People Beware? Investigating Smartphone Versus Nonsmartphone Dependency Symptoms Among Young Singaporeans. *Social Behavior and Personality: An International Journal*. https://doi.org/10.2224/sbp.2015.43.7.1209

Lin, Y. H., Chiang, C. L., Lin, P. H., Chang, L. R., Ko, C. H., Lee, Y. H., & Lin, S. H. (2016). Proposed diagnostic criteria for Smartphone addiction. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0163010

Marty-Dugas, J., & Ralph, B. C. W. (2018). The Relation Between Smartphone Use and Everyday Inattention. *Psychology of Consciousness: Theory, Research, and Practice*, *5*, 46–64. https://doi.org/10.1037/cns0000131

McCrae, R. R., & Costa, P. T. (1987). Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.52.1.81

Miller, G. (2012). The Smartphone Psychology Manifesto. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691612441215

Mou, D. (2016). Battling severe mental illnesses with smartphones: how patients' smartphone data cna help improve clinical care. *MHealth*, *2*.

Oliver, E. (2010). The challenges in large-scale smartphone user studies. In *Proceedings of the 2nd ACM International Workshop on Hot Topics in Planet-scale Measurement - HotPlanet '10*. https://doi.org/10.1145/1834616.1834623

Panova, T., & Carbonell, X. (2018). Is smartphone addiction really an addiction? *Journal of Behavioral Addictions*. https://doi.org/10.1556/2006.7.2018.49

Parker, W. D., & Stumpf, H. (1998). A validation of the five-factor model of personality in academically talented youth across observers and instruments. *Personality and Individual Differences*. https://doi.org/10.1016/S0191-8869(98)00016-6

Piwek, L., Ellis, D. A., & Andrews, S. (2016). Can programming frameworks bring smartphones into the mainstream of psychological science? *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.01252

Przybylski, A. K., & Weinstein, N. (2017). A Large-Scale Test of the Goldilocks Hypothesis: Quantifying the Relations Between Digital-Screen Use and the Mental Well-Being of Adolescents. *Psychological Science*, *28*(2), 204–215. https://doi.org/10.1177/0956797616678438

Richardson, M., Hussain, Z., & Griffiths, M. D. (2018). Problematic smartphone use, nature connectedness, and anxiety. *Journal of Behavioral Addictions*. https://doi.org/10.1556/2006.7.2018.10

Rosen, L. D., Lim, A. F., Felt, J., Carrier, L. M., Cheever, N. A., Lara-Ruiz, J. M., … Rokkum, J. (2014). Media and technology use predicts ill-being among children, preteens and teenagers independent of the negative health impacts of exercise and eating habits. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2014.01.036

Rosen, L. D., Whaling, K., Carrier, L. M., Cheever, N. A., & Rokkum, J. (2013). The Media and Technology Usage and Attitudes Scale: An empirical investigation. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2013.06.006

Rozgonjuk, D., Levine, J. C., Hall, B. J., & Elhai, J. D. (2018). The association between problematic smartphone use, depression and anxiety symptom severity, and objectively measured smartphone use over one week. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2018.05.019

Shaw, H., Ellis, D. A., Kendrick, L.-R., Ziegler, F., & Wiseman, R. (2016). Predicting Smartphone Operating System from Personality and Individual Differences. *Cyberpsychology, Behavior, and Social Networking*, *19*(12), 727–732. https://doi.org/10.1089/cyber.2016.0324

Shaw, H., Ellis, D. A., & Ziegler, F. V. (2018). The Technology Integration Model (TIM). Predicting the continued use of technology. *Computers in Human Behavior*, *83*, 204–214. https://doi.org/10.1016/j.chb.2018.02.001

Sivadas, E., & Venkatesh, R. (1995). An Examination of Individual and Object-Specific Influences on the Extended Self and its Relation to Attachment and Satisfaction. *Advances in Consumer Research*.

Takao, M., Takahashi, S., & Kitamura, M. (2009). Addictive Personality and Problematic Mobile Phone Use. *CyberPsychology & Behavior*. https://doi.org/10.1089/cpb.2009.0022

Tao, S., Wu, X., Zhang, S., Tong, S., Hao, J., & Tao, F. (2017). Association of alcohol use with problematic mobile phone use and depressive symptoms among college students in Anhui, China. *Journal of Public Health (Germany)*. https://doi.org/10.1007/s10389-016-0766-z

Tran, D. (2016). Classifying Nomophobia as Smart-Phone Addiction Disorder. *UC Merced Undergraduate Research Journal*. https://doi.org/10.1111/desc.12173

Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2017). Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among

U.S. Adolescents After 2010 and Links to Increased New Media Screen Time. *Clinical Psychological Science*, *6*(1), 3–17. https://doi.org/10.1177/2167702617723376

UK Parliament. (2018). Impact of social media and screen-use on young people's health inquiry launched. Retrieved August 24, 2018, from https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/news-parliament-2017/social-media--young-peoples-health-inquiry-launch-17-19/

Wang, T., Brede, M., Ianni, A., & Mentzakis, E. (2018). Social interactions in online eating disorder communities: A network perspective. *PLoS ONE*, *13*(7), 1–17. https://doi.org/10.1371/journal.pone.0200800

Wilcockson, T. D. W., Ellis, D. A., & Shaw, H. (2018). Determining Typical Smartphone Usage: What Data Do We Need? *Cyberpsychology, Behavior, and Social Networking*, *21*(6), 395–399. https://doi.org/10.1089/cyber.2017.0652

Wolniewicz, C. A., Tiamiyu, M. F., Weeks, J. W., & Elhai, J. D. (2018). Problematic smartphone use and relations with negative affect, fear of missing out, and fear of negative and positive evaluation. *Psychiatry Research*. https://doi.org/10.1016/j.psychres.2017.09.058

World Health Organization (WHO). (2018). ICD-11 International Classification of Diseases 11th Revision. Retrieved March 3, 2019, from https://icd.who.int/

Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., … Steinberg, D. (2008). *Top 10 algorithms in data mining*. *Knowledge and Information Systems* (Vol. 14). https://doi.org/10.1007/s10115-007-0114-2

Yildirim, C., & Correia, A. P. (2015). Exploring the dimensions of nomophobia: Development and validation of a self-reported questionnaire. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2015.02.059

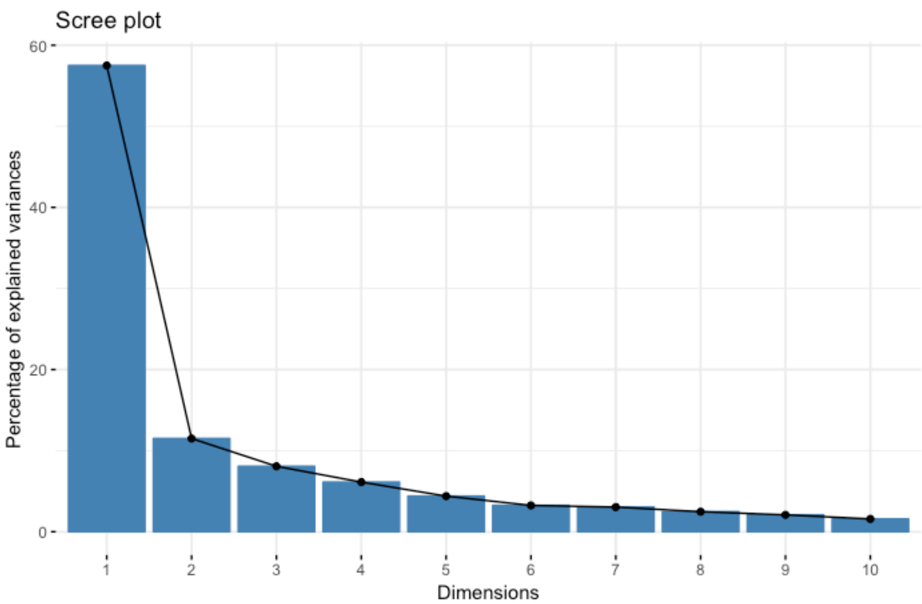
Zhao, S., Ramos, J., Tao, J., Jiang, Z., Li, S., Wu, Z., … Dey, A. K. (2016). Discovering Different Kinds of Smartphone Users Through Their Application Usage Behaviors. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 498–509). Heidleberg, Germany. https://doi.org/10.1145/2971648.2971696

Appendix

Appendix 1. Principal Components Analysis on Smartphone 'Usage' Scales

Due to the scales utilized in this chapter struggling to capture simple usage behaviors and did not improve even when combined, I conducted a Principle Components Analysis (PCA) in order to investigate whether this data could be reduced. PCA is traditionally used to reduce the dimensionality of big data while containing most of the information in the dataset in order to make it more easily analyzed (Brems, 2017; Jaadi, 2019). Interestingly, when I ran a PCA over the scales gathered in *Chapter II*, it is clear that each of the scales clearly load well onto **one** component/dimension (Figure A). This therefore suggests that these scales are not measuring different types of usage as the names of these scales suggest (e.g., nomophobia, smartphone addiction scale, incorporation into the extended self, etc.), but rather they are all measuring an extremely similar construct.

Figure A. A Scree Plot based on the scales used in Chapter II. This plot represents the percentage of variances explained by each principal component. There is a large drop in variance explained moving from one component to two components, where one would only use one component moving forward with any analysis.



Additionally, I looked at the eigenvalues of the scales (dimensions). Table A below shows the eigenvalues, the percentage of variance explained in the dataset, and the cumulative variance explained in the dataset. The typical rule-of-thumb is to keep dimensions that have an eigenvalue of >1 (Abdi & Williams, 2010). Therefore, here, we would keep two dimensions within the dataset as the second dimension just meets the requirements. However, the additional percentage of variance explained may not be sufficient to justify the additional dimension, which is explored by visualizing the biplot of these dimensions (Figure B).

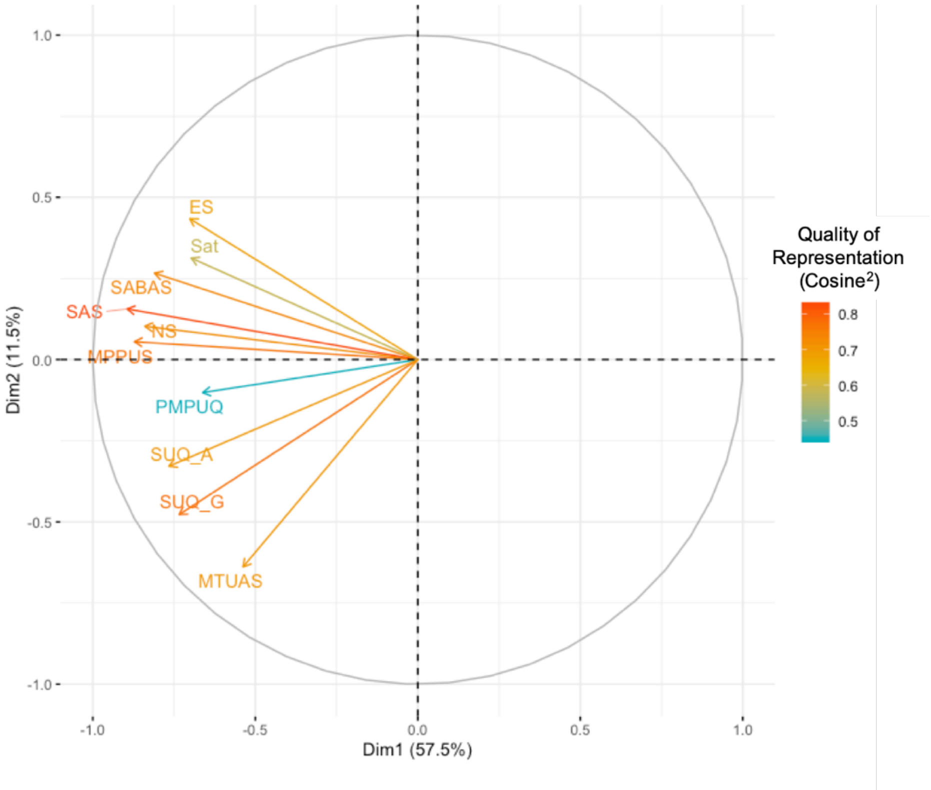
Table A. Eigenvalues, the percentage of variance explained in the dataset, and the cumulative variance explained in the dataset.

	Eigenvalue	**Variance Explained (%)**	**Cumulative Variance Explained (%)**
Dim.1	*5.75*	*57.50*	*57.50*
Dim.2	*1.15*	*11.50*	*69.00*
Dim.3	0.81	8.07	77.07
Dim.4	0.61	6.12	83.19
Dim.5	0.44	4.39	87.58
Dim.6	0.32	3.25	90.83
Dim.7	0.30	3.04	93.87
Dim.8	0.25	2.47	96.34
Dim.9	0.21	2.08	98.42
Dim.10	0.16	1.58	100.00

Finally, we can visualize these two dimensions (Figure B) and assess the quality of representation across all of the scales. This reveals which scale(s) contribute the most or are the most influential to the dimensions. In the figure below, we can see that most scales are aligned to dimension one (57.5% variance explained) (x-axis) and very few are aligned with dimension two (y-axis) (11.5% variance explained), which provides additional evidence for one dimension. Additionally, we can see the quality of representation (cosine^2), which demonstrates which scale(s) are contributing most to the dimensions. This is denoted by the color, with red being the highest influence and blue being the lowest.

Here, we note that the Smartphone Addiction Scale (SAS) (Kwon et al., 2013) is the most representative. Finally, it is worth noting that the smaller the angle between each of these scales denotes the strength of correlation between these scales, also shown by the scales grouping into one cluster (Gabriel, 1971) in Figure B. Further adding to these results, these scales are all highly correlated, which indicates they are measuring very similar constructs.

Figure B. Plot showing the first two dimensions of the PCA. x-axis is dimension 1, y-axis is dimension two. The arrows are each of the scales, which all point towards the left-hand quadrants, which indicate alignment with dimension 1. The coloring of the arrows denotes the quality of representation, whereby red denotes higher quality and blue denotes lower quality.



This demonstrates, again, that these typical 'usage' scales are unable to capture different behaviors, despite their claims to. Hence, reiterating points made in both *Chapter I* and *II*, we need to create and use reliable and accurate measures. This analysis will be a part of a wider set of technology and health studies.

CHAPTER III

THE EVOLUTION OF ONLINE COMMUNITIES

Chapter publication status:

Davidson, B. I., Jones, S. L., Joinson, A. N., & Hinds, J. (2019). The evolution of online communities. *PLoS ONE*. 14(5): e0216932. 10.1371/journal.pone.0216932.

Statement of Authorship (to preface each co-authored paper)

<table>
<tr><th colspan="10">This declaration concerns the article entitled:</th></tr>
<tr><td colspan="10">The evolution of online ideological communities</td></tr>
<tr><th colspan="10">Publication status (tick one)</th></tr>
<tr><td>Draft manuscript</td><td></td><td>Submitted</td><td></td><td>In review</td><td></td><td>Accepted</td><td></td><td>Published</td><td>X</td></tr>
<tr><td>Publication details (reference)</td><td colspan="9">Davidson, B.I., Jones, S. J., Joinson, A. N., & Hinds, J., (2019), The evolution of online ideological communities, PLoS ONE. 14(5): e0216932. 10.1371/journal.pone.0216932</td></tr>
<tr><th colspan="10">Copyright status (tick appropriate statement)</th></tr>
<tr><td colspan="3">I hold the copyright for this material</td><td>X</td><td colspan="5">Copyright is retained by the publisher, but I have been given permission to replicate the material here</td><td></td></tr>
<tr><td>Candidate's contribution to the paper (detailed, and also given as a percentage).</td><td colspan="9">Design of methodology:
BID [33], SJ [33], AJ [33]

Experimental work:
BID [70] conducted all the data analysis with SJ [30].

Presentation of data in journal format:
BID [90] wrote the manuscript and revisions of the manuscript. AJ, JH, and SJ edited each iteration [10].</td></tr>
<tr><td>Statement from Candidate</td><td colspan="9">This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.</td></tr>
<tr><td>Signed</td><td colspan="5">Brittany Davidson</td><td colspan="2">Date</td><td colspan="2">27th June 2019</td></tr>
</table>

Chapter II focused on the traditional tools created to measure technology 'usage' within psychological science (Ellis, 2019; Ellis, Davidson, Shaw, & Geyer, 2019). While the use of self-report measurement is a typical method within psychology, this chapter showed that these 'usage' scales cannot reliably be used as a proxy to measure actual behavior and interaction with devices (e.g., Andrews, Ellis, Shaw, & Piwek, 2015; Ellis, 2019; Ellis et al., 2019; Orben & Przybylski, 2019). Hence, moving forward, new data collection methods will need to be employed to accurately capture and assess any form of technology usage behavior. This is also because if people cannot correctly report simple usage behaviors, like time spent on their smartphone, it is seemingly less likely they could report on usage and interactions with individual devices, apps, or systems. Additionally, as shown in *Chapter* II, many technology usage scales are extremely similar conceptually (*Chapter 2, Appendix 1*), which also leads to a question as to what construct these scales are actually capturing (Davidson & Ellis, 2019; Ellis, 2019).

However, one way to handle these issues is to utilize 'actual', real-world behavior from individual's interactions with technologies, also known as digital footprints or traces (Weaver & Gahegan, 2007). As discussed briefly in *Chapter II*, smartphones and other systems (e.g., reddit, other social media networks) allow for 'in situ' data collection or there are methods to 'scrape' data online (Ellis et al., 2019). For instance, as shown in the previous chapter, we gathered data via Apple's Screen Time. However, there are increasing numbers of other methods to collect various types of data using externally built apps, for example, 'Funf in a box' for screen time measurement (Andrews et al., 2015) or PegLog for location-tracking (Geyer, Ellis, & Piwek, 2018). Similarly, various types of data can be scraped from websites, communities, or social media sites (e.g., metadata and content data). Both of these data types can offer much insight about the user (e.g., Angeletou, Rowe, & Alani, 2011; Pfeil, Svangstu, Ang, & Zaphiris, 2011; Skowron, Ferwerda, Tkalčič, & Schedl, 2016; Welser et al., 2011).

Here, this chapter addresses both overarching research questions of this thesis, where it utilizes time series data to examine whether users do indeed adapt their behavior over time, as well as providing a novel method and approach to understanding user behavior over time. This chapter analyzed metadata from two online ideological communities: RevLeft and Islamic Awakening (N = 1631; N = 849 at time of data collection). The aim was to examine changes in user behavior in terms of their roles within the community and whether users do change roles over time. Additionally, we ground this work theoretically via social role theory and also a leadership framework in order to understand the hierarchy of roles in these communities, and the potential pathways are to leadership.

Online communities provide users with rich sources of information, the ability to exchange ideas, and the opportunity to form social connections (Bateman, Gray, & Butler, 2011). Online communities may be a place for positivity, for instance, users finding support (e.g., illness), meeting new people with shared interests, or sharing artwork or music (Fullwood et al., 2019; Obst & Stafurik, 2010; Ren, Kraut, & Kiesler, 2007; Wang, Brede, Ianni, & Mentzakis, 2018). However, they may have the opposite effect, for example, (specific areas of) 4Chan are highly malevolent, making itself known as the "*internet hate machine*" (Bernstein et al., 2011, p. 51). 4Chan has been attributed to a number of controversies online, varying from hacking politicians in the US, to hoax celebrity deaths, massacre threats, posting nude photographs of celebrities, and images of murder victims (Alfonso & Bond, 2017).

This demonstrates how different online communities have a unique set of users, behaviors, goals, and motivations for use, which makes them an interesting place to study online. Understanding user behavior within online communities (good or bad) is important for community managers, where this will help with moderation of communities. Further, it has the potential to enhance marketing and advertising campaigns as marketing managers could identify influential people within the community in order to disseminate products and services to a wider audience, or potentially security practitioners to identify users leading and guiding narratives within communities associated with crime or terrorism.

While research has looked into roles adopted by users online (Fisher, Smith, & Welser, 2006; Gleave, Welser, Lento, & Smith, 2009; Welser et al., 2011), we found that most often, these roles were viewed and analyzed as a static pattern of behavior, such that once a role is established, it does not tend to change. Understanding that the roles we have offline change and evolve over time, for instance, one may be a daughter, a semi-professional athlete, a student, and a bartender part-time – we must remember that these roles may not necessarily be the same in the next few months, yet alone in the years to come. Hence, the next paper seeks to account for the dynamic nature of roles and focus on how we can understand how users change their roles over time by identifying their behavioral patterns via meta-data over time.

This article first proposes a novel method to analyze behavioral meta-data from online sites to identify various user roles. It is grounded by Preece and Schneiderman's Reader-to-Leader Framework (2009), which provides a theoretical way to understand these roles in terms of leadership and engagement within the community. We argue this is an important contribution to the literature and various contexts (e.g., marketing, security, community moderation) as it allows us to identify leaders (and various other types of user) within communities. Secondly, this paper analyzed roles over a two-year period, where we were able to build a model showing user role transitions over time. This has shown common role transition pathways, less common ones, and sheds light on user turnovers within the communities analyzed, which is an important metric for those managing online communities.

Abstract

Understanding user behavior is valuable to organizations and has applications from marketing to security, for instance, identifying leaders within a community or predicting future behavior. There remain unanswered questions regarding online communities, and we seek to understand the various roles users adopt in online communities – for instance, who leads the conversation? Who are the supporters? We examine user role changes over time and the pathways users follow, which allows us to explore the differences between users who progress to leadership positions and users who fails to develop influence. We also reflect on how user role proportions impact the overall health of the community. We examine two online ideological communities, RevLeft and Islamic Awakening (N = 1631; N = 849). We provide a novel approach to identify various types of users. Further, we studied user role trajectories over time, and identified community "leaders" from meta-data alone.

Study One examined both communities using K-MEANS cluster analysis of behavioral meta-data and revealed eight types of user role. We then mapped these roles against Preece and Schneiderman's Reader-to-Leader Framework (RtLF) (Preece & Schneiderman, 2009). Both communities aligned with the RtLF, where most users were "contributors", many were "collaborators" and few were "leaders". Study Two looked at one community over a two-year period and found despite a high churn rate of users, roles were stable over time. We built a model of user role transitions over the two years. This can be used to predict user role changes in future, which will have implications for community managers and security focused contexts (e.g., analyzing behavioral meta-data from forums and websites with known to be associated with criminal activity).

1 Introduction

Online communities provide users with rich sources of information, the ability to exchange ideas, expertise, and the opportunity to form social connections (Bateman et al., 2011). This can be a source for good, for instance, research has revealed the positive effects of online communities for support (Fullwood et al., 2019; Obst & Stafurik, 2010; Ren et al., 2007; Wang et al., 2018). However, this is not always the case. There are malevolent online communities, for example, (specific areas of) 4chan, also described as the "*internet hate machine*" (Bernstein et al., 2011, p. 51). Other examples include the recent "involuntary celibate", or "incel" movement, found in subsections of 4chan, Reddit, other websites (that may have had links to the Canadian terror attack in 2018 (Beauchamp, 2018; Williams, 2018)), or specific ideologically motivated forums with malicious intent and the aim to radicalize members (Aly, Macdonald, Jarvis, & Chen, 2017).

Users in online communities are largely transient in nature, which tends to cause a high churn rate within these communities (e.g., Bateman et al., 2011; Ransbotham & Kane, 2011) and complicates the understanding of user behavior due to a lack of a consistent user-base. However, despite high churn rates, there tends to be a core set of members that continue to contribute, share, and retain knowledge within a community (Ransbotham & Kane, 2011), which is important for new joiners. While most users will inevitably leave the community, there is often an influx of new users that serves to refresh and maintain membership levels.

Despite this, relatively little is known about how (some) users evolve into valued members once they have become engaged within online communities. Previous research has addressed specific stages in the lifecycle of a community member. For instance, how users join and become accepted (e.g., Dabbish, Farzan, Kraut, & Postmes, 2012; Yuqing et al., 2012; Zhu, Kraut, & Kittur, 2014), or how language is used by those in leadership positions (e.g., Huffaker, 2010). Various researchers have also examined the roles that a user may adopt in a community (Fisher et al., 2006; Gleave et al., 2009; Welser et al., 2011), however this work tends to treat a role as a pattern of behavior that is static (rather than dynamic), such that once a role is established, it rarely changes. This is perhaps a limiting assumption, in light of research that has found that behavioral patterns and roles change over time, both on- and offline (Fiske, 2010; e.g., Gleave et al., 2009; Sheldon, Ryan, Rawsthorne, &

Ilardi, 1997). We adopt various roles throughout our lifetime – changing subtly and substantially according to the context and those around us at the time (e.g., Fiske, 2010).

In the present research, we analyze the roles that users adopt within two moderately-sized online communities; one political discussion forum, RevLeft (denoted as community A) (N > 1000 at each six-month time slice over a period of two years) and one religious discussion forum, Islamic Awakening, denoted as community B (N = 849). Since data collection, both online communities have closed. Community A was a far-left forum with many groups and threads relating to anarchy and various forms of communism (Internet Archive Wayback Machine, 2016). Community B described itself as a "*small effort and a humble contribution […] towards the global revival of Islam*" (Internet Archive Wayback Machine, 2010).

In the first instance (Study One), these roles are treated as stable across the tenure of the community members in order to build a classifier. This classifier is then used to categorize members of one community in unique six-month time slices (Study Two), which allows us to study any movement between roles across time. We also consider the stability in user numbers per role, as an indicator of community heath, as suggested in the work of Angeletou et al. (2011).

We also analyze user churn in community A. Rate of churn is a long-standing concern for community managers who attempt to retain users (Bateman et al., 2011; Ma & Agarwal, 2007; Soroka & Rafaeli, 2006). We analyze all role transitions over the two-year period (N = 7,712) in order to understand the distribution of users' changing roles, remaining in the same role, or potentially leaving the community.

We ground our work within a framework conceptualizing user engagement and leadership: the Reader-to-Leader Framework (RtLF) (Preece & Schneiderman, 2009), which describes how users transition from being passive "readers" to potentially active community "leaders". Hence, the RtLF provides us with a valuable theoretical lens through which we examine user behavior within online communities, focusing on user role changes over time. By grounding our analyses within this framework, we discuss how this is important for those moderating and managing online communities. Finally, we consider the role transition pathways that users make over time and provide a method and theoretical underpinning to understand these various pathways.

2 Theoretical Background

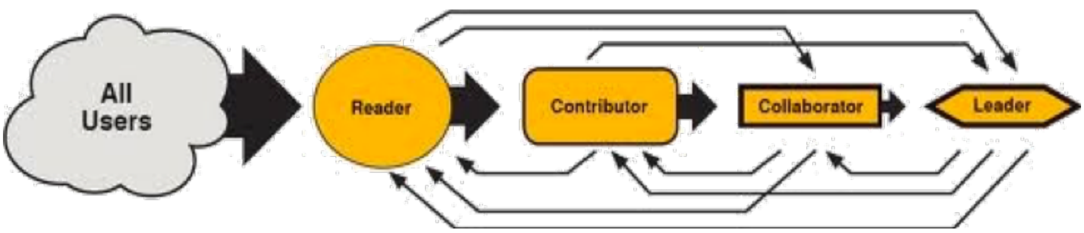
2.1 Conceptualizations of User Engagement

Engaged, active users are the lifeblood of a virtual community. It should be of little surprise then that there have been numerous attempts to study not only user engagement, but also the ways in which users move from passive consumers to active creators within a community. One relatively well established approach is the Reader-to-Leader Framework (RtLF) (Preece & Schneiderman, 2009) (Figure 1). The RtL framework describes four roles that users can adopt in online communities:

- **Reader** – visiting, reading, searching, returning
- **Contributor** – posting, reviewing, rating
- **Collaborator** – engaging with other members, collaborating to create content
- **Leader** – mentoring new joiners, setting policies and monitoring users, promoting participation

Preece and Schneiderman (2009) state that although these categorizations are not exhaustive, they describe the participation of many users. While the RtLF does not specify quantities of users at each stage in the framework, they explicitly state that the proportion of users moving towards a leadership significantly diminishes.

Figure 1. Reader-to-Leader Framework (RtLF)



This inequality in participation has been noted before – for instance, Nielsen (2006) described the 90-9-1 rule of online participation, stating that 90% of users online will lurk – where they "*read or observe, but do not contribute*" (Nielsen, 2006), 9% will have small contributions over time, and the final 1% drive the majority of conversation. Both of these conceptualizations attempt to describe the wide variance of online user participation. The RtLF has been extensively used since its initial publication in 2009 and remains influential when examining online user behavior (Arazy et al., 2017; Panteli, 2016). While it does describe how users can progress from being a reader of content, to a contributor and collaborator sharing their own content, to potentially a leader guiding narratives within the community, it lacks clear criteria for what behaviors constitute each role, and how transitions actually happen. The framework suggests most users follow a linear progression through the successive levels, with a decreasing proportion of users moving from one role to the next (illustrated by the size of arrows in Figure 1). However, it also suggests that users can move in a non-linear fashion (Gilbert, 2017). For example, a portion of users might be able to make the direct transition to a position of leadership, having previously contributed very little to the community. Moreover, the RtLF framework does not offer a strong indication of the proportion of users that make such transitions, nor does it shed light on the characteristics of users that follow particular pathways of participation. The present study addresses these issues by examining users over time with reference to the RtLF in order to quantify proportions of users who progress in both linear and non-linear pathways. While it is useful to understand general patterns in users' paths through their usage lifecycle, we contend that it is also useful to understand the specific roles and trajectories of certain individuals. We propose that understanding user roles will aid understanding of subtle differences in groups of users, which will then offer insight into the dynamic within the community. Such insights could enable us to understand what makes a user maintain, increase, decrease, withdraw participation, and recognize factors that differentiate users that follow different paths through the various levels. Forecasting the future actions of users given their past and present trajectories is likely to be useful for analyzing the health of online communities, and for enabling designers and managers to identify characters such as "rising stars" or "fading leaders" at an early stage.

2.2 Understanding Changes in Engagement

An extensive body of literature examines the different roles that people assume within online communities (Ang & Zaphiris, 2010; Arazy et al., 2017; Panteli, 2016; Welser et al., 2011). In particular, *social role theory* considers behavior to be the enactment of socially defined categories (e.g., teacher, student, manager) (Fiske, 2010). A social role consists of norms, expectations, and behaviors that a person tends to fulfill.

Gleave et al. (2009, p. 1) define social roles as, "*a combination of social psychological, social structural, and behavioral attributes*". Social role theory implies that in order to change behavior (e.g., increased participation) it would be necessary to change roles. Therefore, one might hypothesize that role changes act as an indicator of transitions within the RtLF.

It has been argued that each user possesses a set of beliefs, which may or may not align with the community or group beliefs online or in offline groups. However, as users integrate and interact with a community, they will enter a process of adopting, adapting, and potentially discarding prior beliefs of roles. Hence, every community (on- or offline) is unique and varies in terms of members, behavior, and communications (Chan, Hayes, & Daly, 2010; Ellinas, Allan, & Johansson, 2017). Schmader and Sedikides (2018) proposed a conceptual framework, State Authenticity as Fit to the Environment (SAFE), which provides an additional way to understand engagement. In SAFE, if the individual has a good self-concept, goal, and social fit to the environment, they are more likely to approach and engage.

A number of different roles have been identified in studies of online discussion groups (see Table 1). These roles have primarily been identified through ethnographic study of the content of interactions (Donath, 1998; Marcoccia, 2004), although some effort has been made to use behavioral metrics to recognize these roles (Turner, Smith, Fisher, & Welser, 2005; Viegas & Smith, 2004). The roles that emerged from these studies have shown various levels of depth in terms of a) how specific a role is, b) how dynamic a role is, and c) how dynamic the network is. Furthermore, changes in roles are often dismissed in the analysis.

Table 1. Examples of Roles Identified in Previous Research

Author(s)	**Roles Identified**
Golder & Donath (2004b)	Newbie, Celebrity, Lurker, Flamer, Troll, Ranter
Turner, Smith, Fisher, & Welser (2005)	Answer person, Questioner, Troll, Spammer, Binary poster, Flame warrior, Conversationalist
Campbell, Fletcher, and Greenhill (2009)	Big man, Sorcerer, Trickster
Chan, Hayes & Daly (2010)	Joining conversationalist, Grunt, Taciturn, Popular participants, Popular initiator, Supporter, Ignored

Pfeil, Svangstu, Ang, & Zaphiris (2011)	Moderating supporter, Central supporter, Active member, Passive member, Technical expert, Visitor
Welser, Lin, Cosley, Dokshin, Smith, Kossinets, & Gay (2011)	Substantive experts, Technical editors, Counter vandalism contributors, Social networkers
Panteli (2016)	Emergent leaders, Appointed leaders, Community founder, Sustaining leaders
Arazy, Lifshitz-Assaf, Nov, Daxenberger, Balestra, & Coye (2017)	Role-Article Samplers, Role Embracing, Article Embracing, Role-Article Polymathing

Both Gleave et al. (2009) and Welser et al. (2011) visualized these systematic patterns of behavior as forms of "structural signatures", which provide insight into the overall role distributions within a community. Further, the topologies of communities rely on user individuality in terms of their behavior and levels of participation in order to group them into roles. This then could be used to provide insight into the health of an online community (Angeletou et al., 2011). Our work aims to improve the classification of community members by considering various behavioral metrics (see Table 2). Chan et al. (2010) used nine features to profile users, including popularity, reciprocity, length of interaction, initiation, neighbor's roles, and volume of communication measures, which is a similar set of features utilized in the present studies. Arazy et al. (2017) examined role-transitions within Wikipedia specifically, where they categorized users into four-types of role changes, which sheds light on the fact that user behavior is indeed active and dynamic. Similarly, Campbell, Fletcher, and Greenhill (2009) explore users changing roles within communities, however, they focus more specifically on conflict between roles within communities.

We extend this work by finding social roles in two online communities (Study One: community A and B) and by examining these roles over time (Study Two: community A). First, we analyze the user role changes over time to investigate the most common role transitions for users, for example, how often do users become leaders, and do those leaders often fall from grace? Then, by analyzing more specific role transitions, we are able to understand the proportions of users who engage in various role transitions. This analysis reveals what transitions are more common, which could act as a basis to predict how a user may transition in future.

2.3 Community Health

The relationship between the health of an online community and user churn has two competing schools of thought. The first, and perhaps more traditional perspective, attempts to understand how community managers can increase user engagement and retain as many users as possible, as high churn rates are deemed as negative (Arguello et al., 2006; Bateman et al., 2011; Ma & Agarwal, 2007). However, one might argue that user churn is a natural occurrence of online communities and one should rely more on whether the community is growing overall in size, as opposed to focusing on users becoming inactive only. The second approach argues that the high churn rate is actually a positive trait for an online community. For instance, Soroka and Rafaeli (2006) suggest that if all users (including readers or lurkers) were actively engaging in the community, this could dilute the knowledge and create unnecessary "*noise*" (e.g., off-topic discussion or too much content), which can be destructive. Further, if all users were constantly posting – who is listening? Hence, a high number of reading users (or "lurkers") is not necessarily a negative trait in an online community (Angeletou et al., 2011; Edelmann, 2016). Further, Angeletou et al. (2011) suggest that higher numbers of elite or popular users are a sign of a healthy community, alongside other indicators such as: having a stable distribution of user roles over time; having a mixture of roles within a community; and lower levels of "ignored" and "low engagement" users. However, Chan, Hayes, and Daly's (2010) findings demonstrate a variety of different role compositions in various online communities. Therefore, factors in community health perhaps stretch beyond role composition. Further, other work has shown that social purity (e.g., sharing the same political view) is important within social networks (Dehghani et al., 2016), which suggests that community health is indeed multi-faceted and unique to each community.

The present research will investigate various ways to analyze community health by considering user churn over time, number of roles found, and the stability of them over time.

Our research questions are as follows:

1. What specific roles can be identified using meta-data from online communities?

2. Do user roles change over time? What are the most common pathways and transitions for users?
3. What can we learn about the health of the community based on user roles?

The present research contains two studies. Study One utilizes the meta-data from users within communities A and B, where we used a clustering algorithm (K-MEANS) in order to detect groups of similar users. After analyzing and naming each cluster, we then map the roles established to the RtLF (Preece & Schneiderman, 2009) in order to conceptualize differing levels of engagement, hierarchy, and leadership within the communities. Study Two is a time-series analysis of community A over a two-year period. Here, we are able to understand whether a user changes their behavior to the extent that their new behavioral pattern exemplifies a different role. Further, we examine the distribution of clusters over time with reference to the RtLF, which provides a theoretical grounding of the types of users that participate in communities. We also analyzed every user role change over the two-year period (N = 7,712) and used this to build a model of role transitions presented in Figures 6 and 7.

3 Data Preprocessing

3.1 Data Collection

Content and meta-data were collected from two publicly accessible communities denoted: community A (RevLeft) and community B (Islamic Awakening). This data was collected via "screen scraping". This is a technique that is similar to automated cut-and-pasting from online webpages. This was done by a custom PERL/MySQL tool, which collected data securely from the Internet utilizing a Privoxy/TOR chain. Where needed for forum access, cookies were supplied with HTTP requests made by the tool and were regularly rotated to ensure anonymity. Data errors were captured by validating the number of fields of each type that had been identified and extracted by the HTML parser, and regular expressions from each webpage. All validation errors, each URL scraped, and cookie and IP rotations have been logged and retained in order to monitor scraping accuracy. No scraping behind logins was conducted and only publicly available data was collected and stored in a MySQL database, which complies with the Terms of Service at the time of data collection. User ID's were encrypted via the MD5 hash algorithm to ensure user identity and privacy.

When these forums were scraped, community A had approximately 1.49 million posts and 11,778 active members. Community B had approximately 500,000 posts and 3,205 active members. Both communities A and B have since been closed.

The two present studies only use six months (Study One – A and B) and two years (Study Two – A only) worth of these data archives. This was due to the early years of the archive capturing the start of the forums, therefore, we focused only on the most recent data (at the point of data collection) to analyze as this is when the forums were fully established. For Study One, we used the most recent six-months' worth of data compiled into one data frame. Study Two used two years' worth of data, which was split into four six-month time slices, to allow for a time-series analysis. While having six-month time slices is a broad window within which to classify users into a single behavior-based role, we selected this time period in order to avoid capturing minor or temporary fluctuations in behavior, as opposed to sustained changes in behavior as reflected as changes in roles.

3.2 Data Metric Development

From all scraped data, we derived several types of behavior metrics seen in Table 2. As we intend to group similar users into clusters, we needed to develop metrics in order to compare users against each emerging role, and the community (Angeletou et al., 2011). Typically, when examining user online metadata, prior work had emphasized a variety of features that should be included. For example, Chan, Hayes, and Daly (2010) created structural features (providing an indication of communication between unique users), reciprocity features (how much users reply to one another), popularity features (number of in-neighbors e.g., those who replied to that user, or thanks rate), persistence features (indicates length of conversation and number of places online they post, e.g., in multiple threads or subforums), and initialization features (how often does a user start a thread?). Angeletou et al (2011) also utilized similar metrics overall, which align with the features in the present studies. Hence, prior work has demonstrated the importance of the metrics used in order to distinguish between roles that users adopt in online communities.

Table 2. Metrics Used in Cluster Analysis for communities A and B

Structural Features	**In-Degree**	Total number of unique network neighbors replying to (or quoting) a user
	Out-Degree	Total number of unique network neighbors receiving posts from (or being quoted by) a user
Content Features	**Word Count**	Mean average word count for all of a user's posts
	Percentage Question Marks	Percentage of a user's posts that contain question marks (excluding within URLs)
	Percentage URLs	Percentage of a user's posts that contain URLs
Popularity Features	**Thank Rate**	Mean average number of thanks per post. Calculated as: Total Number of Thanks Received / Total Number of Posts Made
Initiation Features	**Initiation Ratio**	Calculated as: Number of threads initiated/Number of threads participated in
Diversity Features	**No. Threads**	Total number of threads participated in
	No. Sub Forums	Total number of sub-forums participated in
Persistence Features	**Posts Per Sub Forum**	Calculated as: Total number of posts/Number of sub forums participated in
	Posts Per Thread	Calculated as: Total Number of posts/number of threads participated in
Reciprocity Features	**Percentage Bi-directional Neighbors**	Calculated as: Number of neighbors that a user has both received posts from and posted replies to/Total number of unique network neighbors

We also wanted to understand as much as possible about the content posted by users, from a meta-data perspective, hence we added additional "Content Features", which consist of average word count, number of URLs present in posts, and the number of questions asked per post. We anticipated that this would help to distinguish between various low engagement users (or contributors referring to the RtLF). For example, Golder and Donath (2004b) identified "newbies", who tend to have little knowledge but would ask many questions. Therefore, users with a high percentage of questions asked and relatively low numbers of posts might fall into this category. Similarly, we wanted to capture the number of URLs being provided in posts, as we anticipate that this could reflect users signposting information, which might be similar to Welser, Gleave, Fisher, and Smith's "Answer Person" (2007). We also included "Diversity" features, which aims to capture the extent to which a user posts in the same threads or subforums. We believe this could be insightful, particularly for distinguishing between users who focus on a limited number of specific threads, and those who engage in a number of conversations across the broader community.

3.3 Statement of Ethics

The present study was reviewed and approved by the Ethical Implications of Research Activity (EIRA) process within the University of Bath, School of Management (reference number: 2393). IRB approval was not required for this work as it only utilized secondary data analysis.

4 Study One: Cluster Analysis

This study utilizes all user activity during the six-month period prior to data collection. Our aim is to understand sets of similar users within both forums, based on the important behavioral features of each role. This will reveal which roles tend to lead and influence the community, who might support the leaders, and who simply follows. These roles will be mapped against Preece and Schneiderman's (2009) RtLF based on key features of each role, reputation scores ("*the opinions of all your forum users*", based on "*how [their] posts are scored by other forum participants*" (vBulletin Solutions, n.d.), which is a native, inbuilt metric), and the number of users within each role.

4.1 Methods & Results

4.1.1 Metrics Utilized for Analysis

We performed a K-MEANS cluster analysis using Weka 3.8.2. We used this method as it is widely used for behavioral analytics, data mining, and data science more generally (Bernstein, et al., 2011; Arazy et al., 2017; Chan et al., 2010; Jackson et al., 2016; Wu et al., 2008). It also provides easily understandable and scalable outputs (Wu et al., 2008). However, we note that K-MEANS clustering algorithms, while widely applicable, can be sensitive depending on initial seeding (Yedla, Pathakota, & Srinivasa, 2010), therefore additional attention to cluster centers and the variable input is critical.

Table 2 shows the metrics that were used during the clustering process. We also collected each user's "Reputation" score, which was removed from the cluster analysis. Instead, we used this as an additional metric to map the outputs from the cluster analysis to the RtLF, since we expect reputation scores to increase as users progress through the RtLF.

4.1.2 Ideal Number of Clusters

Cluster analyses are considered a form of unsupervised learning due to a lack of a defined set of classes prior to learning (Pan, Shen, & Liu, 2013). K-MEANS is a partitional clustering algorithm that divides the data into smaller sections called "clusters" (Bholowalia & Kumar, 2014). When running the K-MEANS algorithm a pre-defined number of clusters, *k*, is required. The ideal *k* value can be found via trial and error and is highly subjective (Hamerly & Elkan, 2004). We found *k* via the Elbow Method, see Figure 2. This aims to visualize and explain the "*percentage of variance explained as a function of the number of clusters*" (Bholowalia & Kumar, 2014, p. 18). This means that the first few clusters will have large decreases in variance as each additional cluster continues to add information to the model. This will eventually plateau as the model does not continue to improve substantially (Bholowalia & Kumar, 2014). We highlighted the boundaries of the potential number of *k* in Figure 2.

Figure 2. Elbow Plot or "Sum of Squared Errors" Plot for communities A (A) and B (B). The dotted red lines denote the upper and lower boundaries for the ideal number of clusters, k.

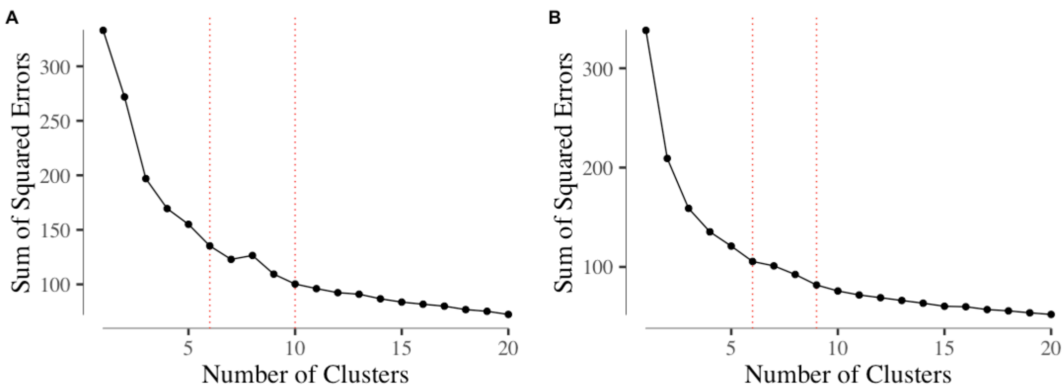


Figure 2 above shows the Elbow Plots for communities A and B. Based on the elbow plot, we proceeded with seven clusters for both communities (using the same *k* value for each community as a basis for comparing how the overall composition of roles differs between communities). We note that in Figure2(A), at $k = 8$, the elbow began to rise, showing that $k = 7$ for this community is optimal. Similarly, in Figure 2(B), the elbow begins to plateau from $k = 7$, which indicates this would be an appropriate number of clusters. Hamerly and Elkan (2004) stress the difficulty of identifying *k*, where they state this tends to rely on prior knowledge. From our theoretical underpinning of the RtLF, we could only take forward $k \geq 3$ (reflecting the contributor, collaborator, leader distinction), however, this would be too coarse for revealing more subtle roles within each of these levels, which have been identified in previous work. In the literature referenced in Table 1, the number of roles identified varies from 3 to 8. We wanted to capture the more intricate and subsets of roles within the RtLF, hence, we proceeded with $k = 7$ as this seemed most appropriate based on Figure 2, previous literature (e.g., Chan et al., 2010; Golder & Donath, 2004a; Turner et al., 2005), and the RtLF.

4.2 Clusters & Visualization

In order to map our clusters to the RtLF, we first need to understand what each cluster means. From looking at the centroids (or multi-dimensional mean) of each variable within each cluster, we were able to deduce how clusters differ from one another. Tables 3 (community A) and 4 (community B) describe each cluster. While we use the same names for the eight roles found in both communities, there are some subtle differences. For example, the Popular users in community B had the highest overall *thanks rate*, whereas the Elite users had the highest *thanks rate* in community A.

Table 3. Cluster Descriptions for community A. See Table 4. for more detailed information on cluster centers.

Cluster/Role Name	Description
Newbie	High initiation rate, highest overall number of questions asked in the community, and typically long word counts in their posts. Lowest in- and out-degree and number of posts
Popular Supporter	High overall metrics, particularly in- and out- degree, bi-directional neighbor degree, and thanks rate. Users were

	similar to the Elite users, however, overall lower in each metric
Taciturn	Low in all metrics, largely not engaged, as reflected by their low activity (e.g., low number of posts and connectivity)
Conversationalist	High number of posts per thread and initiation rates, with high bi-directional neighbors. Low in most other metrics, particularly thanks rate
Elite	Highest in- and out-degree, thanks rate, number of posts, posts per subforum, and typically had posts with low word counts
Low Volume Supporter	Typically, moderate in all metrics, however, often a high number of questions per post and long posts
Information Provider	Longest posts by a substantial measure with the highest number of URL links and a high initiation rate. Moderate in most other metrics

Table 4. Cluster Centers for community A. Red to green coloring indicates the lowest to highest values per metric (row).

Input Variable	Overall	Newbie	Popular Supporters	Taciturn	Conversat-ionalist	Elite	Low Volume Supporter	Information Provider
In Degree	40.68	6.26	83.43	8.75	9.22	234.83	14.96	11.58
Out Degree	42.41	3.97	87.38	8.76	8.11	255.16	14.46	10.08
Total Posts	78.52	6.40	133.74	10.00	15.96	610.97	18.12	22.32
Mean Word Count	107.76	157.44	91.42	75.39	107.01	98.90	115.44	279.35
Thank Rate	0.71	0.60	1.01	0.57	0.50	1.13	0.59	0.74
% Questions per Post	0.29	0.77	0.28	0.07	0.22	0.30	0.40	0.22
% URLs per Post	0.07	0.03	0.06	0.03	0.03	0.06	0.05	0.66
Mean Posts Per Thread	2.06	1.91	2.53	1.34	4.78	2.65	1.86	1.78
Initiation Ratio	0.21	0.55	0.09	0.16	0.56	0.08	0.17	0.37
Mean Posts Per Subforum	8.56	2.94	16.25	2.51	8.00	37.77	4.31	4.72
% Bi- directional Neighbors	0.21	0.12	0.32	0.09	0.59	0.45	0.18	0.18

Table 5. Cluster Descriptions for community B. See Table 6. for more detailed information on cluster centers.

Cluster/Role Name	**Description**
Elite	Highest in- and out-degree, thanks rate, number of posts, and posts per subforum, and typically short posts.
Newbie	Highest overall number of questions asked in the community, and typically long word counts in their posts. Low in all other metrics.
Low Volume Supporter	Moderately low in all metrics, however, a high initiation rate.
Popular Supporter	High overall metrics, particularly mean posts per subforum, in-, out-, and bi-directional neighbor degrees. Low initiation ratio and URLs in posts. Similar to Elite, however, overall lower in each metric.
Conversationalist	Highest initiation rate, high word counts, and bi-directional neighbors. Low in- and out-degrees, and thanks rate.
Taciturn	Lowest in all metrics, aside from slightly higher connectivity (in- and out-degree).
Information Provider	Longest posts, highest number of URLs per post and initiation rate. Lowest in all other metrics.

Table 6. Cluster Centers for community B. Red to green coloring indicates the lowest to highest values per metric (row).

Cluster Input	ALL	Elite	Newbie	Low Volume Supporter	Popular Supporter	Conversat-ionalist	Taciturn	Information Provider
In Degree	28.66	149.05	8.24	12.96	45.37	5.38	8.61	1.95
Out Degree	29.44	158.89	8.32	10.97	46.90	3.65	9.16	0.87
Total Posts	72.24	489.44	12.75	22.74	92.15	13.64	11.74	5.23
Mean Word Count	118.04	96.00	150.16	118.02	116.49	163.88	67.08	217.03
Thank Rate	0.62	0.89	0.66	0.61	0.64	0.50	0.63	0.19
% Questions per Post	0.29	0.30	0.67	0.23	0.30	0.28	0.10	0.28
% URLs per Post	0.08	0.03	0.03	0.08	0.04	0.06	0.03	0.77
Mean Posts Per Thread	2.24	2.50	1.85	1.73	3.79	2.27	1.41	1.25
Initiation Ratio	0.28	0.13	0.12	0.44	0.10	0.95	0.04	0.77
Mean Posts Per Subforum	10.06	39.90	3.90	5.87	15.44	6.66	3.46	3.60
% Bi-directional Neighbors	0.26	0.51	0.19	0.19	0.42	0.31	0.12	0.09

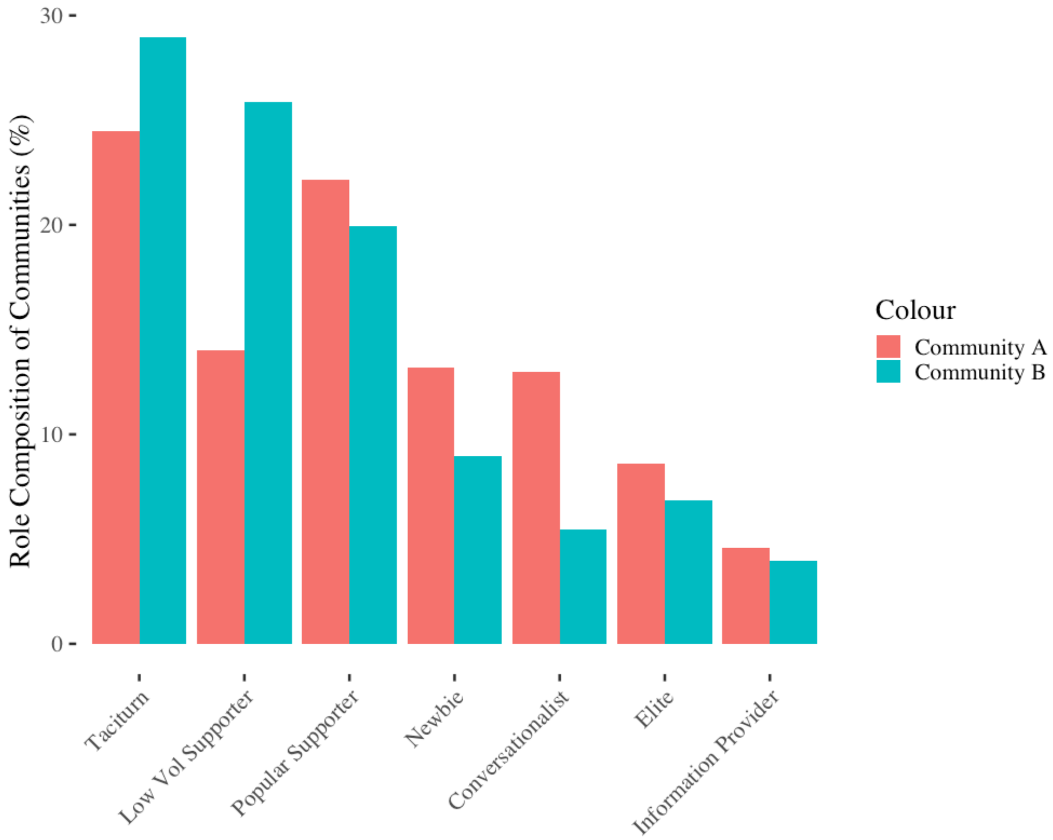
4.3 Mapping Clusters to the Reader-to-Leader Framework

Each role will belong to the leader, collaborator, or contributor categories from the RtLF. Readers were not directly included, as they were originally described as users who are "*venturing in, reading, browsing, searching, returning*", (Preece & Schneiderman, 2009, p. 16), hence they have no engagement and passive behavior without a digital trace. These users may not have created an account until they became a contributor, collaborator, or leader, as both forums at the time of data collection were open to view and browse.

We assessed the similarity of clusters identified for each community. While, there are subtle variances in each role identified, we found that they are similar enough to map to the RtLF in the same way. For example, the Popular Supporters in community A and B were subtly different. Community A's Popular Supporters had high thanks rates, whereas this was not as reflected in community B. However, both community A and B's Popular Supporters had high in- and out- degrees, bi-directional neighbors, with high mean posts per thread and subforums. Both community's Taciturns were low in all metrics, however, community B's tended to be more connected (e.g., higher in- and out-degree scores). Similarly, both community's Elite users had much in common (e.g., high in- and out- degrees, bi-directional neighbors, thanks rates, and number of posts). However, community B's number of URLs was a much lower value for Elite users. This highlights subtle differences between forums. This is to be expected, in light of social role theory and literature that addresses the dynamic between the individual and group identity, and shows that there is negotiation and changes in user behavior and adoption of (new) beliefs as individuals integrate into groups and as communities evolve (e.g., Ellinas et al., 2017). Further, this aligns with the work of Chan, Hayes, and Daly (2010) where they demonstrated the unique role compositions of different forums.

Next, we examined the proportion of users in each of the roles for both community A and community B (Figure 3), for comparison against the RtLF, as well Nielsen's 90-9-1 rule of social media and online community engagement (2006). We anticipated seeing larger numbers of users in roles such as Low Volume Supporter or Taciturn, in comparison to Elite users, which was accurate for both forums.

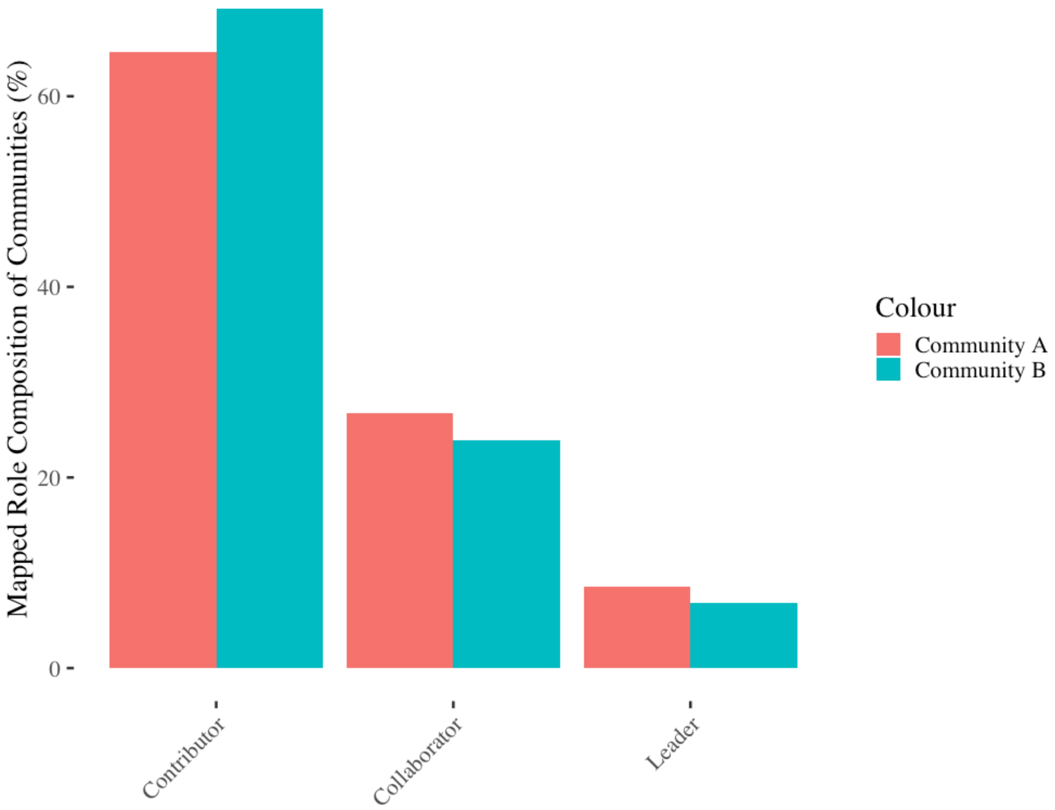
Figure 3. Proportion of Users in each Cluster for community A and community B



Communities A and B were similar in terms of numbers of users in each role. For instance, in both platforms the highest number of users fell into the roles of Low Volume Popular Supporters, Taciturns, and Newbies. Interestingly, community A had a higher proportion of users falling into the Conversationalist role than community B, and community B had a higher proportion of Low Volume Supporting users. This further demonstrates the subtle differences in role composition in online communities (Chan et al., 2010; Ellinas et al., 2017). Preece and Schneiderman (2009) suggest that reputation is associated with roles, where leaders would have the highest reputation, followed by collaborators with a moderate reputation, and contributors with lower reputation. Hence, the collected reputation scores (inbuilt metrics within the community forum software) were used to further inform this stage. The reputation score for each user is calculated as a function of their total number of posts and reputational upvotes or downvotes from other community members (which are weighted by the reputation *power* that other users wield). However, we acknowledge that the precise details for how the reputation metric is calculated is unknown, which is a key limitation. Hence, we place higher importance on the key features of each cluster and the number of users within each cluster, then we consider reputation as an additional guideline. Reputation was often helpful with our conceptualized metrics, for example, in both forums the reputation score of Elite users far exceeded that of all other clusters.

We noted that the Information Providers and Popular Supporters in community B had lower reputation scores than expected. With consideration of their key features (e.g., in- and out-degree, thanks rates, longer posts), we decided that the Information Providers and Popular Supporters fulfilled the criteria to be a "collaborator" ("*developing relationships, working together, setting goals*" (Preece & Schneiderman, 2009, p. 20)). Figure 4 shows the proportion of users belonging to clusters mapped to either contributor, collaborator, or leader, for communities A and B (Figure4).

Figure 4. Proportion of Users in Each Category of the Reader-to-Leader Framework (RtLF) for A (A) and B (B)



Both forums had high numbers of contributors, moderate number of collaborators, and few leaders, which aligns with the RtLF. Despite differences in the proportion of users in the unmapped roles (Figure 3), when mapped to the RtFL, these differences are reduced. Community A's distribution of users as contributors, collaborators, and leaders was 69.22%, 23.91%, and 6.87%, respectively, whereas community B's users were 64.66% contributors, 26.74% collaborators, and 8.60% leaders.

5 Study One Discussion

Study One aimed to first identify roles using behavioral meta-data. Second, we analyzed these roles by framing them in terms of the RtLF (Preece & Schneiderman, 2009). From our scraped meta-data, we found seven clusters in communities A and B. They ranged from low engagement and passive users, such as Taciturns, to Information Providers, with high levels of engagement, to Elite users, with the highest levels of popularity and thanks rates across the community.

This study shows that various types of role and levels of engagement can be detected within online communities via cluster analysis. These roles can be used to better understand the social structure within online communities. Preece and Schneiderman (2009) state their framework is not exhaustive, however, it is dynamic and captures the majority of user behavior online. The contribution of this research is a deeper exploration of contributors, collaborators, and leaders, alongside the identification of "sub-roles" that sit within each category from the RtLF. Further, seeing user types at a higher resolution than the RtLF provides deeper insight into the subtle dynamics within a community. For instance, there were differences in frequency of each individual role between communities A and B, however, these were less noticeable and could be missed if we had only considered three roles: contributors, collaborators, and leaders. We note there were subtle differences in specific role categories, which is to be expected across online communities. Our findings align with Chan, Hayes, and Daly (2010), showing that forums are unique communities with different compositions of roles. A further explanation for these variations could be the size differences of the forums, as B (N = 849) is approximately half the size of A (N = 1,631). However, a key limitation of this study is that we were limited to meta-data only. It may be possible to shed further light on the subtle differences between online communities by also analyzing the linguistic content. Further, we acknowledge methodological limitations, such as the use of K-MEANS clustering algorithm. Due to the dependence between several metrics, we utilized K-MEANS as no assumptions would be violated, specifically the lack of independence in our variables. The trade-off is the potential sensitivity of K-MEANS, although we mitigated this by utilizing several methods to find a suitable number of clusters, *k*, and thoroughly checking cluster centers against theory and literature. We believe this was the most suitable clustering algorithm for this study due to its ability to handle high-dimension datasets and its increasing use in behavioral analytics (Arazy et al., 2017; Chan et al., 2010; Jackson et al., 2016; Wang et al., 2018; Wu et al., 2008).

Study One shows that we can detect various types of users from behavioral meta-data alone. We build on this in Study two, where we analyze user behavior from community A over time – whether these roles tend to be stable for the user, if they tend to change, and whether the number of users in each role stays consistent. This has several implications for community managers who moderate online communities, marketers looking to identify potential influencers and endorsers of brands, and there may also be implications for security contexts.

6 Study Two: Role Transitions & Community Health

Study Two focused exclusively on community A. First, we classified users into the roles from Study One for each of the additional three six-month time slices. This allowed us to analyze the stability of users' roles and calculate the most common role transitions that users experienced. This aimed to investigate if users shift over time, and more specifically, if (and how) users became leaders, which is shown in Figures 6 and 7. Here, we are most interested in understanding whether users do indeed change behavior as they continue their engagement with this community. Further, this analysis also revealed community A's high user churn, despite the overall increase in number of users. In addition to the six-month time slice in study one, we used three additional six-month time slices for the period up to the 24 months before data collection date (Table 7). The time slice used in Study One was used as a training set for a Naïve Bayes classifier utilized in Study Two.

Table 7. Each Six-Month Time Slice and Number of Users in community A

Time Period	Number of Users
-24 months to -18 months	1,293
-18 months to -12 months	1,458
-12 months to -6 months	1,495
-6 months to time of data collection	1,631

6.1 Classification of Community A's Users Over Time

During the classification step, we used a Naïve-Bayes algorithm, which is a type of generative classifier (Ng & Jordan, 2001). This works by taking the inputs (here, these were the metrics found in Table 2) and making predictions of the label (here, this would correspond to the clusters revealed in Study One), where the user is assigned to the most likely cluster they belong to (Ng & Jordan, 2001). We classified the final three time slices to the roles detected in the Study One, which allowed us to examine the stability of the roles (mapped and unmapped to the RtLF) over time (Figure 5).

Upon classification, we performed various sensitivity analyses, shown in Tables 8 and 9 below. We note that the sensitivity and accuracy measures show reasonable classification performance. Particularly, as seen in Table 8, the ROC Area for all clusters was, on average, 0.96, which is regarded as "excellent" (Fawcett, 2006). This is further demonstrated by the high true positive (TP) rate (column 1) and the false positive (FP) rate remaining low (column 2).

Table 9 is the confusion matrix from our classification step, which provides more detail regarding which clusters were less accurately classified. The highest area of sensitivity in the classification model often concerned Low Volume Supporters. They were slightly more likely to be misclassified due to the lack of distinctive features (e.g. they lacked particularly high or low metrics for certain behaviors), unlike the other roles. There is also potential for greater sensitivity among the lower contributing roles, as each of their interactions may have more impact on the overall metrics. However, we wanted to keep the low engagement users in our analysis (e.g., low volume supporters) in order to reflect the broad spectrum from non-engagement to high engagement users presented in the RtLF. This also allowed us to capture users who may be more likely to join and leave swiftly, as shown in our analysis of churn within the community (see Figures 7 and 8). Hence, if we removed these users, we would lose subtle sub-roles that are an important section of the community user base. We reiterate that we are unable to capture readers, since they do not leave a digital trace.

Table 8. Classification Accuracy by Cluster (%)

	TP Rate	**FP Rate**	**Precision**	**Recall**	**F Measure**	**MCC**	**ROC Area**	**PRC**
Newbie	0.87	0.06	0.58	0.87	0.69	0.68	0.95	0.87
Popular Supporter	0.87	0.03	0.88	0.87	0.88	0.85	0.98	0.93
Taciturn	0.88	0.05	0.88	0.88	0.88	0.83	0.98	0.96
Conversationalist	0.60	0.01	0.79	0.60	0.68	0.67	0.97	0.73
Elite	0.99	0.01	0.89	0.99	0.94	0.93	1.00	0.96
Low Volume Supporters	0.62	0.07	0.76	0.62	0.68	0.59	0.92	0.78
Information Provider	0.92	0.01	0.77	0.92	0.84	0.84	0.96	0.87
Weighted Avg.	0.80	0.05	0.81	0.80	0.80	0.76	0.96	0.88

Table 9. Confusion Matrix for Classification (%). Actual values as rows; predicted values as columns

Predicted / Actual	Newbie	Pop Sup	Taciturn	Conver.	Elite	Low. Vol. Sup.	Info. Prov.
Newbie	**7.79**	0.00	0.00	0.25	0.00	0.67	0.25
Popular Supporter	0.00	**17.41**	0.12	0.12	0.8	1.35	0.12
Taciturn	0.18	0.61	**25.38**	0.00	0.00	2.51	0.25
Conversationalist	0.55	0.25	0.49	**3.25**	0.00	0.61	0.31
Elite	0.00	0.06	0.00	0.00	**6.81**	0.00	0.00
Low Volume Supporter	4.97	1.41	2.82	0.43	0.00	**16.06**	0.18
Information Providers	0.00	0.06	0.06	0.06	0.06	0.06	**3.68**

Figure 5. Percentage of Users in each Role – Mapped to the RtLF (A) and Unmapped to the RtLF (B)

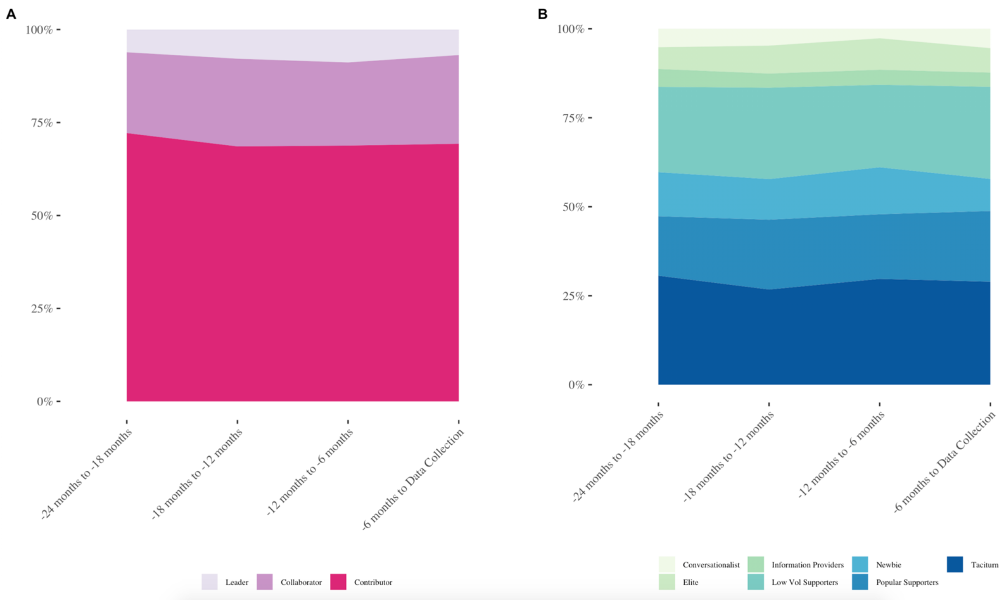


Figure 5 shows that the percentage of users in each role is moderately stable and consistent. We can see this is particularly prominent with the Elite and Information Providers, as well as contributor roles like Low Volume Supporters and Taciturns. We noticed the increase in Low Volume Supporters in the most recent time slice. There are relatively subtle fluctuations in overall RtFL categories. However, in the most recent time slices the numbers of users in each mapped role (contributors, collaborators, and leaders) were remarkably stable, despite the amount of churn in members leaving and joining the community. We found 25.32% of users from the earliest time slice were present at the time of data collection 24 months later. While substantial numbers of the community left at each time slice, the overall size of the community grew over the two-year time period, which reveals that large numbers of new users also joined.

6.2 User Role Pathways

Next, we examined user role changes over the two-year period. We analyzed every possible cluster transition that users could make (N = 64, i.e. switching between clusters, and also transitions to an inactive (reader) state). 7,712 user transitions were observed (i.e., comparisons of an individual's role from one time slice to the next), which included users that remained in the same cluster, changed once, or even multiple times. Users could change up to three times (across the four time slices). Transitions were identified simply by comparing individual user's cluster allocations across consecutive time slices. Only users who had appeared in earlier time slices were included in any subsequent inactive to inactive transitions (i.e., users who had not yet joined the community are not counted).

Table 10 shows the *top ten most common transitions* seen within community A across all time slices. The frequency of each transition pathway was counted, which formed the basis for calculating the most common pathways.

Table 10. Top 10 Most Common Pathways for Users to Take Based on all Possible Transitions (N=7,712)

Pathway (Clusters)	**Number of Users**	**% of Users**
inactive → inactive	1082	14.03
inactive → taciturn	834	10.81
taciturn → inactive	780	10.11
low volume supporter → inactive	629	8.16
inactive → low volume supporter	480	6.22
newbie → inactive	367	4.76
inactive → newbie	335	4.34
popular supporter → popular supporter	324	4.20
inactive → popular supporter	292	3.79
low volume supporter → low volume supporter	195	2.53

Four out of the top ten pathways in Table 10 regard users becoming inactive or remaining inactive. Hence, it was common for users to disappear from the community (or assume a reader role) from one time slice to the next. Four out of the top ten concerned new joiners or those becoming active again, where three of these pathways were users becoming contributors, and the other pathway reflecting users going straight to a collaborator role. The final pathway were users who remained contributors or collaborators.

We developed a model from the all user transitions of community A (Figure 6). This model mimics the RtLF and highlights the diminishing numbers of users progressing to leadership within the community. From Table 10, if we consider the percentage of *all transitions* users made (N = 7,712), only 2.9% of those transitions were of users shifting from a contributor to a collaborator, and only 0.9% of transitions were collaborators transitioning to leaders. Few transitions concerned users joining as leaders (0.9%), however, it was slightly more common to join as collaborators (4.8% of transitions) or simply a contributor (25.3% of transitions).

Figure 6. Model Showing Percentage of All Users Role Transitions Mapped to the Reader-to-Leader Framework

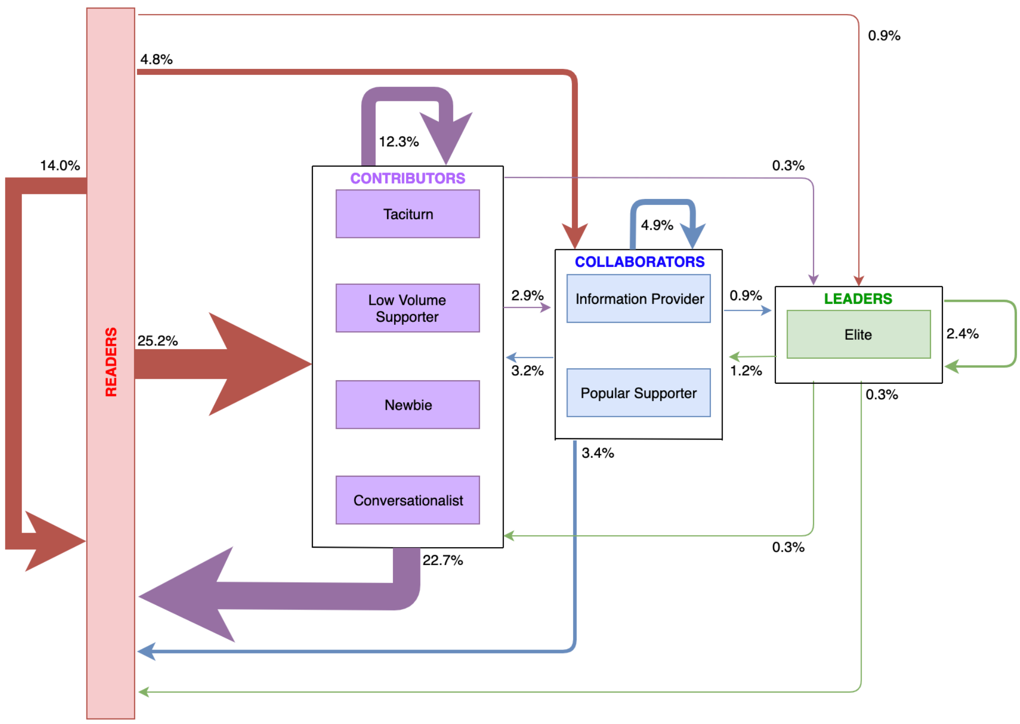
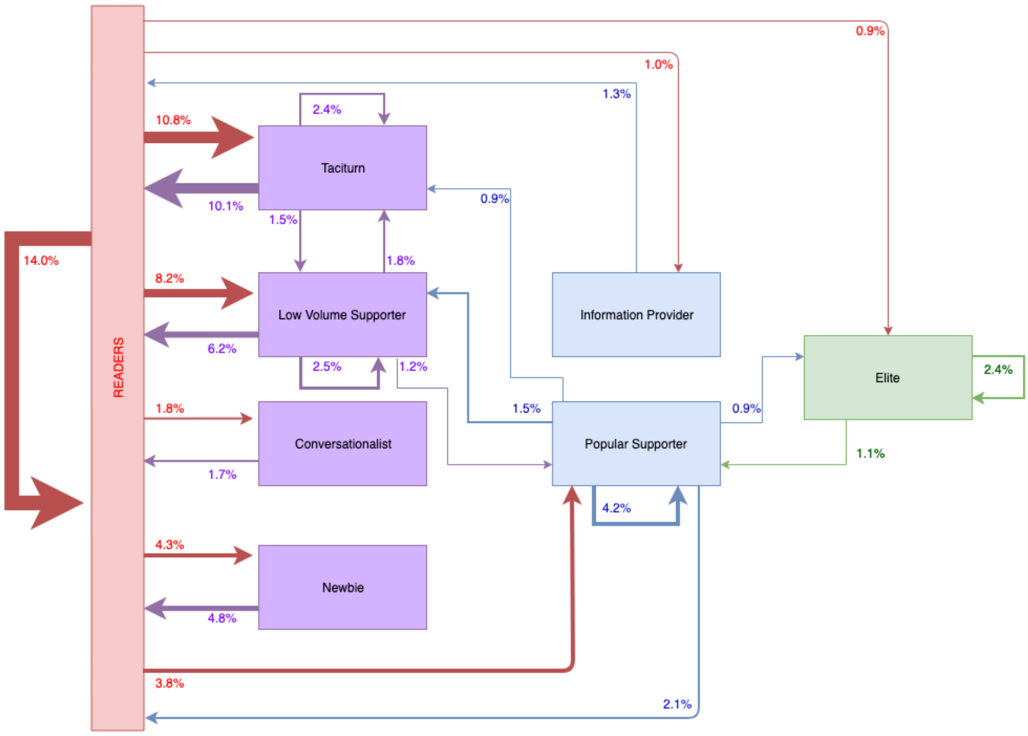


Figure 6 provides a high-level overview of how users changed their roles over time. This considers *all transitions made* in in the dataset. It reveals that users staying in the same RtLF category is common, for instance, we found 12.3% of transitions were users that were once contributors and remain contributors, similarly, 4.9% of transitions were users that did not change from a collaborating role, and 2.4% transitions were users that remained leaders. Figure 6 also demonstrates the high churn of users, where 30.9% of transitions regarded new joiners or those returning from a period of inactivity, whereas 26.4% of the transitions were users changing back to readership/inactivity. This also demonstrates that the community within the two-year time period of data grew overall.

In addition to Figure 6, we present Figure 7, which shows the top 25 transitions users made. This, in contrast to Figure 6, is not mapped against the RtLF, which enables us to see the intricate (role level) pathways users took. It also shows the most common pathway users took to become a leader (although this appears to be exceptionally rare in community A). It further demonstrates non-linear pathways of users (e.g., demoting transitions – Popular Supporter to Taciturn).

Figure 7. Model Showing the Top 25 Transitions of Users Role Transitions. Demonstrating the linear and non-linear pathways users took



From Figure 6, we saw that 25.2% of transitions during the two-year period, transitioned from readership to a contributor. Figure 7 provides a higher resolution of this and shows that the majority of transitions from readers into contributor roles were users becoming Taciturns, followed by Low Volume Supports, then Newbies, and much less commonly, Conversationalists. We also found a fairly common cyclical role pathway of users either remaining Taciturns or Low Volume Supporters or switching between the two. Therefore, users are more likely to switch between these roles or become inactive rather than progress towards leadership.

If we consider the paths of progression towards leadership, out of all transitions made, it was exceptionally uncommon for any contributor (Taciturn, Low Volume Supporter, Conversationalist, or Newbie) to become an Information Provider. Instead, it was more common for a small number of readers to jump straight to this role. Within the *top 25 most common transitions*, the only contributor role that led to a collaborator role was Low Volume Supporters transitioning into Popular Supporters. However, it was actually more common for Popular Supporters to be demoted back to Low Volume Supporters. Out of all transitions made, 0.9% of transitions progressed from Popular Supporters to the leadership role of the Elite. In terms of the top 25 most common transitions, the most common path to the Elite is: **Reader → Low Volume Supporter → Popular Supporter → Elite.** While, there were other pathways identified, these were exceptionally rare routes that users followed in comparison.

7 General Discussion

We presented two studies that reveal insightful information about the composition and dynamics of an online community based on meta-data alone. For instance, uncovering specific behaviors associated with particular subsets of users within a community, known as roles. These roles can be framed within the RtLF, which helps to reveal which users are in leadership positions and which are less engaged. The first study analyzed the meta-data from two ideological communities (A and B). We found seven clusters via K-MEANS cluster analysis, which we analyzed and formulated into different roles within each forum. These roles were mapped against the RtLF, based on the key features of each role, the number of users in each role, and the average reputation score of each role. The findings from both studies have potential to be used to predict user role changes in future.

Further, the method is a key contribution, as it provides a way to reveal and further examine groups of users over time, providing insight into user dynamics within various online communities. This information is useful for a variety of contexts, for instance, targeting marketing campaigns for specific groups of users and across different contexts or identifying potential new influencers for brands. In addition, this could be of interest to security practitioners as this work, and our methods, may provide insight into which users may be leading and guiding narratives, and a way to identify potential users of interest.

7.1 User Role Compositions

When comparing the two forums, we found similar roles, however, there were some specific differences. For instance, the Popular Supporters in community A and B were subtly different. Where community A's Popular Supporters had high thanks rates, this was not as reflected in community B. However, they were similar in all other metrics (e.g., in- and out-degree, bi-directional neighbors). As noted in Figure3, there are differences in the proportion of users in each role for community A and B, where community B had a higher proportion of Low Volume Supporters, however, community A had a higher proportion of Conversationalists. These differences were to be expected, as each online community is its own eco-system consisting of different roles, proportions of roles, and individuals within it (Chan et al., 2010). Further, the system itself can impact the way in which users behave, thus influencing the expressed behaviors of each role (e.g., Levina & Arriaga, 2014).

Once we had described and developed the roles based on features, frequency, and reputation scores, we then mapped these roles to the RtLF. Both forums had high numbers of contributors, many collaborators, and few leaders. However, we were unable to capture "readers", who are "*venturing in, reading, browsing, searching, returning*", (Preece & Schneiderman, 2009, p. 16), as they would not have generated any behavioral meta-data captured via scraping. We analyzed these roles, both mapped (to the RtLF) and unmapped, in community A over a two-year time period (Study Two). This was to firstly provide insight into how stable these roles are over time and what this implies about the health of an online community. Secondly, we considered how users changed over time and which role transitions were most common. Further, this allowed us to examine user churn, which provided further insight into the intricate social dynamic of the community.

Interestingly, we found there was a large churn of users at each time slice (Dabbish et al., 2012; Ransbotham & Kane, 2011). However, as seen in Table 7, the overall size of community A grew at each time slice. Despite this large turnover of users over time, and the constant influx of new users or returning users from a period of inactivity, as seen in Figure 5, the proportion of users falling into each role (mapped or unmapped to the RtLF) was remarkably stable. This has implications for the overall health of an online community (Angeletou et al., 2011), which will be discussed further in the following sections. Focusing specifically on the user role compositions, both mapped and unmapped, based on Figures 3 to 5, we see a calm and stable appearance of user turnover and proportions of users in each role. However, there is a flurry of role transitions taking place beneath the surface.

7.2 The Pathway to Leadership

We revealed users changing roles across each time slice – rarely moving in a linear fashion through the RtLF, but more commonly, staying in the same RtLF category, or demoting their role. We reiterate that Figures 6 and 7 and Table 10 are based on the number of users who made each transition over *all time slices.*

We analyzed every transition each user made (N = 7,712) and discovered that the majority of transitions were users becoming inactive or users who remained inactive (40.4%). This aligns with literature looking at user participation, motivation, and the retention of users (e.g., Bateman et al., 2011; Ma & Agarwal, 2007; Soroka & Rafaeli, 2006). It is often difficult to retain users and motivate them to engage (Ma & Agarwal, 2007), which was seen within community A, where the churn of users was high. However, we discovered that the second most common set of transitions were new users joining the community into contributor and collaborator positions, which aligns with our findings that despite the high churn of users, the overall community grew in size over the two-year time period. Other relatively common pathways include a cycle between Taciturns and Low Volume Supporters, which consists of users who showed extremely low engagement as Taciturns and somewhat higher levels of engagement as Low Volume Supporters (Table 4). Typically, users in these roles (as contributors more generally) either remained in a contributing role (Figure 6) or became inactive, as shown in Figure 7. These non-linear movements are to be expected with online communities, as reflected in the RtLF (Preece & Schneiderman, 2009) and in Nielsen's Rule of Internet engagement (2006) with the vast majority of users engaging and contributing little.

Perhaps one of our most important empirical findings relates to the rarity of users becoming leaders within community A. We demonstrate that astonishingly few users did progress linearly through the RtLF and following this pathway to the Elite role is against the odds. However, the exceptional users who were leaders, often stayed leaders. In contrast, contributors were far more transient in nature, shown by increased numbers transitioning to inactivity from those roles. However, based on

Figure 7, we did discover the most common pathway users would take to become a leader:

Reader → Low Volume Supporter → Popular Supporter → Leader

Out of all transitions that users made (N = 7,712) in the two-year period, only 1.2% of transitions were from Low Volume Supporters to Popular Supporters, and 0.9% of all transitions involved progression to Elite from Popular Supporter roles. There were other pathways users could take to become leaders, however, they were extremely rare (and were subsequently not captured in Figure 7). This is perhaps not surprising, as other work has shown that recruitment and mentoring of new editors in online communities such as Wikipedia remains a challenge (Musicant, Ren, Johnson, & Riedl, 2011). Hence, we suggest there may be a similar lack of mentorship from Elite users, which contributes to the few users becoming leaders in community A. However, this would need to be investigated further using content data.

7.3 User Roles and Community Health

Within community A, we found that roles, both mapped and unmapped to the RtLF, remained consistent over time. It may be that a healthy online community has a level of stability and consistency of role distributions over time, aligning with Angeletou et al., (2011). If we consider consistent roles alone as a sign of health within a community, we would argue that community A is an example of a healthy community (it only closed recently due to financial reasons - several years after the data was collected for the present research). Angeletou et al. (2011) also noted that increasing levels of "ignored" users will decrease the overall health of a community. Examples of "ignored" users in our analysis were Taciturns, due to their low engagement and overall contribution. We found the numbers of Taciturns remained stable over time, further indicating a healthy community. We also align with Soroka and Rafaeli (2006), where they propose lurking behavior is unlikely to be harmful, as these users may not have content to contribute. They argue that enabling readership without contribution is helpful to maintaining a healthy community, as it can reduce noise and clutter across forums. If we consider that elite users are there to guide and influence the community, too many users attempting to do this could lead to detrimental effects on the community.

We must recognize that there are differences for what constitutes “healthy”. For instance, the specific nature of a community, as types of roles, and the composition of those roles within the community will naturally vary (e.g., Chan et al., 2010). This provides a potential avenue for further research. In the present study, we have focused on ideological communities. We anticipate there could be differences in non-ideological communities due to the nature of the content shared or the purpose of use as demonstrated by Chan, Hayes, and Daly (2010).

7.4 Limitations

First, referring back to Table 2 and the metrics employed, it is important to understand how these relate to behavior presented by users. While these metrics are ego-centric, some features (e.g., structural and reciprocity) also rely on the other community members for feedback (e.g., thanks rate). Features like this are critically important for understanding leading users, as we would anticipate these users to be popular and well-regarded. However, we must note that each community has its own eco-system of roles, which will naturally have different behavioral features (Chan et al., 2010; Ellinas, Allan, & Johansson, 2016), therefore, the optimal metrics to identify particular subsets of users with particular behavioral patterns might differ from community to community. The key limitation here is that, despite the metrics developed for this work, meta-data can only provide a certain level of information. Analyzing the forum content data or performing a network analysis, for example, would likely reveal further insight about the intricacies of specific online communities.

Second, we acknowledge that the overall category of “contributor” is wide in our dataset, where we have included users with an extremely low levels of engagement (e.g., <10 posts). This can cause difficulty with classification, as noted in Table 9, where Low Volume Supporters were the most likely to be misclassified. However, as stated previously, these users were kept in the dataset as we aimed to capture the entire community, especially those closer to the uncaptured “readers”. This is an important section of the community to include, as these may be users that have just joined, or are close to leaving the community. We also note that extremely low activity users could create highly sensitive metric values, which may have an impact on the accuracy of the clustering values. However, as seen in Tables 8 and 9, where misclassification was low, we do not see strong evidence that this was a significant issue.

Finally, the time slices used (six months) are also wide. We selected this window primarily as we wanted to ensure that changes in behavior reflected significant role changes, rather than capturing temporary fluctuations of engagement. We do acknowledge, however, that some users may have changed multiple times within the six-month time period, which were naturally uncaptured here.

7.5 Future Work

The present study demonstrates that we can examine user behavior to gain understanding of how their role within an online community may change over time. This provides the basis for future research directions. First, one might replicate this across a variety of online communities. This would offer insight into the differences between different types of online communities. This could also utilize more than just meta-data alone (e.g., content, linguistic features), perhaps employing qualitative or ethnographic approaches to reveal other subtleties.

Second, there is work to be done regarding metric development from meta-data (and other data types). Our metrics have been based on literature (e.g., Chan et al., 2010), as well as adding some additional features. We note that due to low activity users being of interest, there are potentially better ways to handle measurements on various engagement metrics such that the overall sensitivity is reduced. The meta-data we used was derived from public postings, but many other forms of data would be available to forum administrators (e.g. profile updating, post deletion, length and frequency of access) that would be useful in building models of user behavior.

Third, further work may consider the use of different theories to ground the modeling, for example, the use of social identity theory to consider in- and out-group differences between roles or communities. Although, this again, would benefit from utilizing more than just meta-data alone.

Fourth, we acknowledge the time slices we used in this work (six months), are fairly wide. Further work could explore different sizes of time slice and what different time slices could offer in terms of understanding user behavior (e.g., subtle versus substantial changes in behavior, temporary or sustained).

Finally, and perhaps most importantly, our empirical findings have highlighted an area of future work relating to understanding moderation, mentoring, and other potential mechanisms that could be used to foster and develop new users (Musicant et al., 2011). This could reveal ways in which we can create the new leaders of online communities by making rare pathways to leadership more widely known and accessible for new joiners.

8 Conclusions

Online communities have the power for good – to support those in need, to create a shared collective consciousness, and to exchange information and ideas. However, this powerful influence can also be utilized conversely. For example, there has been a recent up rise with the "involuntary celibate" or "incel" movement (Beauchamp, 2018; Williams, 2018). We also face a constant battle with radicalization online (Benigni, Joseph, & Carley, 2017), which remains difficult to understand and intervene. We have presented a novel method to examine roles within an online community, which has utilized the Reader-to-Leader Framework (Preece & Schneiderman, 2009) in order to help conceptualize types of users in terms of leadership. This approach has the potential to be highly valuable in contexts where the role evolution of online forum users needs to be investigated. The demonstrated method can be applied to a variety of online behavioral meta-data and used by researchers and practitioners interested in understanding online communities from a data-driven perspective. This research also has implications for community managers and moderators wishing to assess the health of community by understanding role distributions and dynamics, and security analysts wanting to identify how leadership positions are occupied in malevolent online communities.

References

Alfonso, F., & Bond, J.-M. (2017). The most disturbing controversies in 4chan history. Retrieved September 29, 2018, from https://www.dailydot.com/unclick/4chan-controversies/

Aly, A., Macdonald, S., Jarvis, L., & Chen, T. M. (2017). Introduction to the Special Issue: Terrorist Online Propaganda and Radicalization. *Studies in Conflict & Terrorism*, *40*(1), 1–9.

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS ONE*, *10*(10), 1–9. https://doi.org/10.1371/journal.pone.0139004

Ang, C. S., & Zaphiris, P. (2010). Social Roles of Players in Mmorpg Guilds. *Information, Communication & Society*, *13*(4), 592–614. https://doi.org/10.1080/13691180903266952

Angeletou, S., Rowe, M., & Alani, H. (2011). Modelling and Analysis of User Behaviour in Online Communities. In *In International Semantic Web Conference* (pp. 35–50). Berlin, Germany. https://doi.org/10.1007/978-3-642-25073-6

Arazy, O., Lifshitz-Assaf, H., Nov, O., Daxenberger, J., Balestra, M., & Coye, C. (2017). On the "How" and "Why" of Emergent Role Behaviors in Wikipedia. In *In the proceedings of the Conference on Computer-Supported Cooperative Work Conference (CSCW)* (pp. 2039–2051).

Arguello, J., Butler, B., Joyce, E., Kraut, R., Ling, K. S., & Wang, X. (2006). Talk to me: foundations for successful individual-group interactions in online communities. In *Proceedings of the SIGCHI Conference on Human Factors In Computing Systems* (pp. 959–968). https://doi.org/10.1145/1124772.1124916

Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). The Impact of Community Commitment on Participation in Online Communities. *Information Systems Research*, *22*(4), 841–854. https://doi.org/10.1287/isre.1090.0265

Beauchamp, Z. (2018). Incel, the misogynist ideology that inspired the deadly Toronto attack, explained. Retrieved September 13, 2018, from https://www.vox.com/world/2018/4/25/17277496/incel-toronto-attack-alek-minassian

Benigni, M. C., Joseph, K., & Carley, K. M. (2017). Online extremism and the communities that sustain it: Detecting the ISIS supporting community on Twitter. *PLoS ONE*, *12*(12), 1–23.

https://doi.org/10.1371/journal.pone.0181405

Bernstein, M., Monroy-Hernández, A., Harry, D., André, P., Panovich, K., & Vargas, G. (2011). 4chan and /b/: An Analysis of Anonymity and Ephemerality in a Large Online Community. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, (Coleman), 50–57. https://doi.org/10.1.1.207.9761

Bholowalia, P., & Kumar, A. (2014). EBK-Means : A Clustering Technique based on Elbow Method and K-Means in WSN. *International Journal of Computer Applications*, *105*(9), 17–24. https://doi.org/10.5120/18405-9674

Campbell, J., Fletcher, G., & Greenhill, A. (2009). Conflict and identity shape shifting in an online financial community. *Information Systems Journal*, *19*(5), 461–478. https://doi.org/10.1111/j.1365-2575.2008.00301.x

Chan, J., Hayes, C., & Daly, E. M. (2010). Decomposing Discussion Forums using User Roles. *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*, *112*(988), 215–218. Retrieved from http://aran.library.nuigalway.ie/xmlui/handle/10379/1114

Dabbish, L., Farzan, R., Kraut, R., & Postmes, T. (2012). Fresh Faces in the Crowd: Turnover, Identity, and Commitment in Online Groups. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work - CSCW '12*, 245. https://doi.org/10.1145/2145204.2145243

Davidson, B. I., & Ellis, D. A. (2019). Technological Déjà Vu [Rapid Response]. *BMJ*, *365*. Retrieved from https://www.bmj.com/content/365/bmj.l2171/rr-1

Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., … Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, *145*(3), 366–375. https://doi.org/10.1037/xge0000139

Donath, J. (1998). Identity and Deception in the Virtual Community. In *Communities in Cyberspace* (pp. 27–58). https://doi.org/10.1519/JSC.0b013e3181e4f7a9

Edelmann, N. (2016). What is Lurking? A Literature Review of Research on Lurking. In G. Riva, B. K. Wiederhold, & P. Cipresso (Eds.), *The Psychology of Social Networking: Personal Experience in Online Communities Vol. 1* (pp. 159–170). Warsaw, Poland; Berlin, Germany: Dr Gruyter Open Ltd. https://doi.org/10.1515/9783110473780

Ellinas, C., Allan, N., & Johansson, A. (2016). Project systemic risk: Application examples of a network model. *International Journal of Production Economics*, *182*, 50–62. https://doi.org/10.1016/j.ijpe.2016.08.011

Ellinas, C., Allan, N., & Johansson, A. (2017). Dynamics of organizational culture: Individual beliefs vs. Social conformity. *PLoS ONE*, *12*(6), e0180193. https://doi.org/10.1371/journal.pone.0180193

Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior*, *97*, 60–66. https://doi.org/10.1016/j.chb.2019.03.006

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies: Special Issue on Human Accuracy*. https://doi.org/10.1016/j.ijhcs.2019.05.004

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*, 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Fisher, D., Smith, M., & Welser, H. T. (2006). You Are Who You Talk To: Detecting Roles in Usenet Newsgroups. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06); IEEE*. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1579410

Fiske, S. T. (2010). *Social Beings* (2nd ed.). NJ, USA: Wiley.

Fullwood, C., Chadwick, D., Keep, M., Attrill-Smith, A., Asbury, T., & Kirwan, G. (2019). Lurking towards empowerment: Explaining propensity to engage with online health support groups and its association with positive outcomes. *Computers in Human Behavior*, *90*, 131–140. https://doi.org/10.1016/j.chb.2018.08.037

Geyer, K., Ellis, D. A., & Piwek, L. (2018). A simple location-tracking app for psychological research. *Behaviour Research Methods*. https://doi.org/https://doi.org/10.3758/s13428-018-1164-y

Gilbert, S. (2017). Portraits of participation : Exploring the relationship between social motivators and facets of participation in a Twitter-based community. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2293–2302.

Gleave, E., Welser, H. T., Lento, T. M., & Smith, M. A. (2009). A Conceptual and Operational Definition of ' Social Role ' in Online Community. *Proceedings of the 42nd Hawaii International Conference on System Sciences*, 1–11. https://doi.org/10.1109/HICSS.2009.6

Golder, S. a, & Donath, J. (2004a). Social Roles in Electronic Communities. In *For the Association of Internet Researchers (AoIR) conference Internet Research 5.0* (pp. 1–25).

Golder, S., & Donath, J. (2004b). Social Roles in Electronic Communities. *Internet Research*, *5*, 19–22. Retrieved from papers2://publication/uuid/A44A65B0-3B94-4809-A728-DAB3D4674856

Hamerly, G., & Elkan, C. (2004). Learning the k in k-means. *Advances in Neural Information Processing Systems*, *16*, 281–288.

Huffaker, D. (2010). Dimensions of Leadership and Social Influence in Online Communities. *Human Communication Research*, *36*(4), 593–617. https://doi.org/10.1111/j.1468-2958.2010.01390.x

Internet Archive Wayback Machine. (2010). Islamic Awakening. Retrieved January 29, 2019, from https://web.archive.org/web/20100306072706/http://www.islamicawakening.com:80/helpus.php

Internet Archive Wayback Machine. (2016). Home of the Revolutionary Left. Retrieved January 29, 2019, from https://web.archive.org/web/20160304035019/http:/www.revleft.com/vb/

Jackson, C., Østerlund, C., Maidel, V., Crowston, K., & Mugar, G. (2016). Which Way Did They Go? Newcomer Movement through the Zooniverse. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, 623–634. https://doi.org/10.1145/2818048.2835197

Levina, N., & Arriaga, M. (2014). Distinction and Status Production on User-Generated Content Platforms: Using Bourdieu's Theory of Cultural Production to Understand Social Dynamics in Online Fields. *Information Systems Research*, *25*(3), 468–488.

Ma, M., & Agarwal, R. (2007). Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities. *Information Systems Research*, *18*(1), 42–67. https://doi.org/10.1287/isre.l070.0113

Marcoccia, M. (2004). On-line polylogues: Conversation structure and participation framework in internet newsgroups. *Journal of Pragmatics*, *36*, 115–145. https://doi.org/10.1016/S0378-2166(03)00038-9

McCrae, R. R., & Costa, P. T. (1987). Validation of the Five-Factor Model of Personality Across Instruments and Observers. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.52.1.81

Musicant, D. R., Ren, Y., Johnson, J. A., & Riedl, J. (2011). Mentoring in Wikipedia:

a clash of cultures. In *In Proceedings of the 7th International Symposium on Wikis and Open Collaboration* (pp. 173–182).

Ng, A. Y., & Jordan, M. I. (2001). On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes. In *Neural Information Processing Systems* (pp. 841–848). https://doi.org/10.1007/s11063-008-9088-7

Nielsen, J. (2006). The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities. Retrieved March 21, 2018, from https://www.nngroup.com/articles/participation-inequality/

Obst, P., & Stafurik, J. (2010). Online we are all able bodied: Online psychological sense of community and social support found through membership in disability-specific websites promotes well-being for people living with a physical disabilty. *Journal of Community and Applied Social Psychology*, *20*, 525–531.

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, *3*, 173–182. https://doi.org/10.1038/s41562-018-0506-1

Pan, W., Shen, X., & Liu, B. (2013). Cluster Analysis: Unsupervised Learning via Supervised Learning with a Non-convex Penalty. *Journal of Machine Learning Research*, *14*, 1865–1889. https://doi.org/10.1016/j.cmet.2012.08.002.

Panteli, N. (2016). On leaders' presence: interactions and influences within online communities. *Behaviour and Information Technology*, *35*(6), 490–499. https://doi.org/10.1080/0144929X.2016.1144084

Parker, W. D., & Stumpf, H. (1998). A validation of the five-factor model of personality in academically talented youth across observers and instruments. *Personality and Individual Differences*. https://doi.org/10.1016/S0191-8869(98)00016-6

Pfeil, U., Svangstu, K., Ang, C. S., & Zaphiris, P. (2011). Social Roles in an Online Support Community for Older People. *International Journal of Human-Computer Interaction*, *27*(4), 323–347. https://doi.org/10.1080/10447318.2011.540490

Preece, J., & Schneiderman, B. (2009). The Reader-to-Leader Framework: Motiviating Technology-Mediated Social Participation. *Transactions on Human-Computer Interaction*, *1*(1), 1–25. https://doi.org/10.5121/ijfcst.2014.4403

Ransbotham, S., & Kane, G. K. (2011). Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in

Wikipedia. *MIS Quarterly*, *35*(3), 613–627.

Ren, Y., Kraut, R., & Kiesler, S. (2007). Applying common identity and bond theory to design of online communities. *Organization Studies*, *28*(3), 377–408. https://doi.org/10.1177/0170840607076007

Schmader, T., & Sedikides, C. (2018). State Authenticity as Fit to Environment: The Implications of Social Identity for Fit, Authenticity, and Self-Segregation. *Personality and Social Psychology Review*, *22*(3), 228–259. https://doi.org/10.1177/1088868317734080

Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of Personality and Social Psychology*, *73*(6), 1380–1393. https://doi.org/10.1037/0022-3514.73.6.1380

Skowron, M., Ferwerda, B., Tkalčič, M., & Schedl, M. (2016). Fusing Social Media Cues: Personality Prediction from Twitter and Instagram. *Proceedings of the 25th International Conference Companion on World Wide Web.*, 107–108. https://doi.org/10.1145/2872518.2889368

Soroka, V., & Rafaeli, S. (2006). Invisible participants: how cultural capital relates to lurking behavior. *Proceedings of the 15th International Conference on the World Wide Web*, 163–172. https://doi.org/10.1145/1135777.1135806

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *28*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Turner, T., Smith, M., Fisher, D., & Welser, H. (2005). Picturing Usenet: Mapping Computer-Mediated Collective Action. *Journal of Computer-Mediated Communication*, *10*(4), 00.

vBulletin Solutions. (n.d.). An Introduction fo User Reputation. Retrieved September 12, 2018, from https://www.vbulletin.com/docs/html/main/reputation_intro

Viegas, F. B., & Smith, M. (2004). Newsgroup Crowds and AuthorLines: visualizing the activity of individuals in conversational cyberspaces. In *Proceedings of the 37th Annual Hawaii International Conference on System Sciences* (pp. 1–10). Hawaii, USA. https://doi.org/10.1109/HICSS.2004.1265288

Wang, T., Brede, M., Ianni, A., & Mentzakis, E. (2018). Social interactions in online eating disorder communities: A network perspective. *PLoS ONE*, *13*(7), 1–17. https://doi.org/10.1371/journal.pone.0200800

Weaver, S. D., & Gahegan, M. (2007). Constructing, Visualizing and Analyzing a Digital Footprint. *Geographical Review*, 97(3), 324-350

Welser, H., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. *Proceedings of the 2011 IConference on - IConference '11*, 122–129. https://doi.org/10.1145/1940761.1940778

Welser, H. T., Gleave, E., Fisher, D., & Smith, M. A. (2007). Visualizing the Signatures of Social Roles in Online Discussion Groups. *Journal of Social Structure*, *8*(2), 1–32.

Williams, Z. (2018). 'Raw hatred': why the "incel" movement targets and terrorises women. Retrieved September 13, 2018, from https://www.theguardian.com/world/2018/apr/25/raw-hatred-why-incel-movement-targets-terrorises-women

Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., … Steinberg, D. (2008). *Top 10 algorithms in data mining*. *Knowledge and Information Systems* (Vol. 14). https://doi.org/10.1007/s10115-007-0114-2

Yedla, M., Pathakota, S. R., & Srinivasa, T. M. (2010). Enhancing K-means clustering algorithm with improved initial center. *International Journal of Computer Science and Information Technologies*, *1*(2), 121–125.

Yuqing, R., Harper, M., Drenner, S., Kiesler, S., Terveen, L., Riedl, J., & Kraut, R. (2012). Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds. *MIS Quarterly*, *36*(3), 841–864. https://doi.org/10.1002/smj.666

Zhu, H., Kraut, R. E., & Kittur, A. (2014). The impact of membership overlap on the survival of online communities. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 281–290. https://doi.org/10.1145/2556288.2557213

CHAPTER IV

CONSISTENTLY INCONSISTENT COMMUNICATION IN ONLINE COMMUNITIES

Thank you, Sophie Ruthven ('Fluffy'), Ana Levordashka, Kseniya Stsiampkouskaya, and Mehrnaz Tajmir for word translations for this chapter's covering image ("*identity*" and "*authenticity*").

Statement of Authorship (to preface each co-authored paper)

<table>
<tr><td colspan="10">This declaration concerns the article entitled:</td></tr>
<tr><td colspan="10">Consistently Inconsistent Communication in Online Communities</td></tr>
<tr><td colspan="10">Publication status (tick one)</td></tr>
<tr><td>Draft manuscript</td><td></td><td>Submitted</td><td></td><td>In review</td><td>X</td><td>Accepted</td><td></td><td>Published</td><td></td></tr>
<tr><td>Publication details (reference)</td><td colspan="9">Davidson, B. I., Ellis, D. A., & Joinson, A. N., (submitted), Consistently Inconsistent Communication in Online Communities.</td></tr>
<tr><td colspan="10">Copyright status (tick appropriate statement)</td></tr>
<tr><td colspan="3">I hold the copyright for this material</td><td>X</td><td colspan="5">Copyright is retained by the publisher, but I have been given permission to replicate the material here</td><td></td></tr>
<tr><td>Candidate's contribution to the paper (detailed, and also given as a percentage).</td><td colspan="9">Formulation of ideas:
BID [90] and AJ [10]

Design of methodology:
BID [90], DAE [10]

Experimental work + data visualization:
BID [100]

Presentation of data in journal format:
BID [90] wrote the manuscript. DAE & AJ [10] helped edited later versions.</td></tr>
<tr><td>Statement from Candidate</td><td colspan="9">This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.</td></tr>
<tr><td>Signed</td><td colspan="5">Brittany Davidson</td><td colspan="2">Date</td><td colspan="2">27th June 2019</td></tr>
</table>

Thus far, in *Chapter* II, I have demonstrated that using traditional proxies of technology usage behaviors (e.g., scales or estimates) are inadequate. This has a number of implications relating to work that intends to measure various interactions with technologies, whether that is time spent, with whom and for how long a user interacts with another, or to examine whether and how behavior changes over time within individual systems or devices. Hence, if the research question is indeed concerned with 'behavior', we must measure objective behavior rather than rely on largely unvalidated measurement tools. In *Chapter* III, we used objective meta-data scraped from two moderate-sized online forums in order to derive behavioral metrics to examine user behavior change over time. Additionally, *Chapter III* demonstrated one can infer roles from behavioral metrics at the group level within online communities based on meta-data. This work revealed that much like individuals offline, users can, and do, change their roles over time within online communities. Whether this is to progress towards a leadership position or slowly becoming inactive; it is clear there are many pathways and they are often non-linear. There are many additional questions that could stem from this work, from predicting user role changes, whether the method proposed, and leadership framework works in different online communities (e.g., are there communities without clear leadership?), and similarly, more methodological questions relating to use of particular algorithms, time slice widths, etc. However, what this chapter did demonstrate, is that we can learn about users from meta-data alone, and this can indicate how active they are, who they are engaging with, when, and this can be used to infer their 'status' within the community. The following chapter is now interested in what can content data reveal about individuals in terms of their behavior in different communities.

While the meta-data gathered and utilized in *Chapter III* provided many insights into user behavior, other user generated data can also be used in provide additional insight into individual or group of users of interest (e.g., location, timestamps, content). With the use of a variety of different data sources alongside meta-data, there is an ability to generate a better understanding of user behavior. For instance, within *Chapter III*, we found a subset of users who left the community for at least 6 months and came back with the same username. This is an interesting behavior that could be better explained by understanding the user's recent content (e.g., suggesting they are taking a break from social media). Similarly, elements of content data (explained in the following sections) can be used to predict various other behaviors, for example group dynamics (Gonzales, Hancock, & Pennebaker, 2010; Sharma & Choundhury, 2018), relationship initiation and stability (Ireland et al., 2011), alongside other real-world outcomes (e.g., grades, life expectancy) (Boyd & Pennebaker, 2017).

In contrast to *Chapter III*, this chapter focuses on content data alone (words only; not image, audio, or video). This chapter addresses both of the key research questions of this thesis: firstly, to examine whether users will change their behavior based on their language across online communities. Additionally, this chapter utilizes a different type of data and analytical approach in order to reveal different insights about users and communities online. Here, I employ both text-mining and big data analytics to traditional statistical methods to compare groups (e.g., ANOVA).

The premise of this project is to analyze whether users will adapt their communication behavior in order to match various community's they have posted in. If they do tend to converge their communication style towards the community, this chameleon-like behavior may lend itself to the notion that users sometimes change their social role in different communities. Hence, this chapter investigates groups of users who have all posted across the same communities and examines whether these users converge or diverge their communication style (specifically linguistic style matching, explained in the next sections) towards the community. Of course, I acknowledge there are many other types of words and content that is shared online (e.g., image, video, audio), however, these data types are not analyzed in this thesis. This is further discussed in the limitations section of this chapter.

Abstract

Individuals naturally adapt their behavior across a variety of contexts. One approach to understanding these changes is from the perspective of social roles, where people adapt behavior according to their environment. Here, and for the first time, we examine in two studies how individual's *linguistic style* changes across different sub-groups within a single online community. Using secondary data from Reddit, we develop a measure of *Linguistic Style Matching Shift*, which indicates whether a user is converging or diverging from sub-group specific norms. The results demonstrate that users are flexible across sub-groups within an online community. However, we also observe limits to this flexibility as users' linguistic style did not shift across similar communities, unless there were explicit differences in community moderation. This work demonstrates how an individual's linguistic style can change across contexts, and has a variety of applied implications including from the impact of moderation to applied security contexts.

1 Introduction

People change and adapt their behavior across contexts (e.g., Fiske, 2010; Niederhoffer & Pennebaker, 2002; Papacharissi, 2002). For instance, as we transition from work to home, patterns of behavior and communication will alter accordingly. This is an example of switching *social roles.* Throughout the lifespan people adopt, lose, and shift social roles as we move from being a daughter to student, colleague, partner, and so on (e.g., Fiske, 2010). People are intrinsically social - upon entering different social contexts, people become aware of individuals around us (who are also enacting social roles), and we respond according to the role(s) they are currently being played (Hogg, Terry, & White, 1995).

Social roles are structured patterns of behavior (Ang & Zaphiris, 2010) that have been used to understand meaningful interactions between individuals within a network or system (Welser et al., 2011). They have been shown to be a useful way to analyze behavior across a range of online contexts, including Wikipedia, forums, and online gaming platforms (e.g., Ang & Zaphiris, 2010; Davidson, Jones, Joinson, & Hinds, 2019; Gleave, Welser, Lento, & Smith, 2009; Pfeil, Svangstu, Ang, & Zaphiris, 2011; Turner, Smith, Fisher, & Welser, 2005; Welser et al., 2011; Welser, Gleave, Fisher, & Smith, 2007). Adopting a social role perspective allows us to view the self as dynamic and ever-changing, formed in response to the changing people and environments around us offline and increasingly online. For instance, we wouldn't expect a person to behave, communicate, and present themselves in exactly the same way on a professional account (e.g., LinkedIn) as they would on a social platform (e.g., Instagram, Facebook), a dating platform (e.g., Tinder, Grindr, Her), or even on a more anonymous platform (e.g., Reddit, 4Chan). However, it isn't clear to what extent this assumption is true – do people really change as they move across contexts online, and in what ways? The aim of the present research is to test this assumption by examining users' behavior across diverse sub-groups within an online community.

Online communities offer a place for people to learn, integrate, and bond with others, much like offline communities (e.g., sports clubs, hobbies). They range from large with a wide range of discussion areas (e.g., Reddit, Usenet) to very specific, single-topic focused communities (e.g., World of Warcraft/DOTA (gaming), Ravelry (knitting), etc.). Online communities have multiple challenges: membership tends to be transient by nature, meaning newcomers must be welcomed and socialized; a large proportion of users are ‘lurkers’ or ‘readers’, and behavior needs to be managed to create a ‘safe’ environment for interaction (Matias, 2019; e.g., Preece & Schneiderman, 2009; Ransbotham & Kane, 2011; Ren et al., 2012; Yuqing Ren, Kraut, & Kiesler, 2007). Typically, users are able to come and go as they please, with or without making accounts, depending on the community (e.g., Bateman, Gray, & Butler, 2011; Ransbotham & Kane, 2011). This transient nature of users in online communities will also impact the social norms over time – online communities, as with offline groups, evolve as members’ behavior changes with time (Davidson et al., 2019) or as they move context. This enactment of social roles can be reflected as behavior, self-presentation, and communication patterns (e.g., Fiske, 2010; Welser et al., 2011). Little is known, however, about how a user’s roles change as they move from one online context to another. While research has demonstrated that different communities consist of different sets or compositions of social roles (e.g., Chan, Hayes, & Daly, 2010); the present research is the first study to examine the same users across different sub-groups or sub-communities. Specifically, the present research is interested in patterns of communication, where we seek to understand whether users change their communication patterns in response to the communicative norms of different communities.

1.1 Communication in Online Communities: CAT Theory

Communication Accommodation Theory (CAT) provides a comprehensive approach to understanding the linguistic adjustments individuals make during social interaction. These sometimes subtle adjustments have been argued to indicate attitudes between two discussants, reflecting the social distance between them (Giles & Ogay, 2007). According to CAT, this accommodation can take a number of forms, including the pace of interaction, matching or mimicry of the actual words used, accent, and in linguistic style (Giles, Coupland, & Coupland, 1991; Giles & Ogay, 2007; Muir, 2016). Interlocutors can *converge* or *diverge* in their communication (or a combination of the two), and such movement is reflective of a variety of group, interpersonal and intra-psychic factors (Muir, Joinson, Cotterill, & Dewdney, 2016), including identification with a social group, desire to integrate/ingratiate, power relations, social attraction, and desire to manage one’s impression. CAT suggests that accommodative behavior is driven by the motivation to gain social approval or to express the desire to assimulate into a group (Giles et al., 1991; Muir et al., 2016). Whereas, of course, divergance typically accentuates differences in speech, linguistic style, etc. (Giles & Ogay, 2007). Thus, according to CAT, the degree to which a user will or will not converge is explained by their sense of self-belonging to the community and how salient that group identity is for them (Giles & Ogay, 2007; Matias, 2019; Yuqing Ren et al., 2007; Schmader & Sedikides, 2018).

While research on communication accommodation is relatively mature, less work has examined how this adapts from context to context, and the degree to which the same user will change their communication (language) behavior in different online communities (with differing social norms). We propose that if a user consistently matches the community communication style, this could be argued as a social role shift with the user adapting in order to fit in with the community (Schmader & Sedikides, 2018). Additionally, by examining the same users across multiple sub-groups we can start to establish whether some users shift their communication patterns more consistently. That is, are some users 'chameleon-like' (Jones, Cotterill, Dewdney, Muir, & Joinson, 2012) because they consistently shift their linguistic behavior in response to a particular group norm.

1.2 Linguistic Style Matching

In the present study we use linguistic style (LS) and linguistic style matching (LSM) – calculated using a 'bag of words' approach to computational linguistics (LIWC (Pennebaker, Booth, Boyd, & Francis, 2015)) to study communication accommodation across different communities. Linguistic style is comprised of an individual's use of function words, which typically consists of nine dimensions (see Data Preprocessing section). Over half of the words used during an interaction are made up of function words (e.g., pronouns, articles, negations). These words carry little meaning on their own (Muir et al., 2016; Muir, Joinson, Cotterill, & Dewdney, 2017; Niederhoffer & Pennebaker, 2002), however, the way an individual uses them and the extent to which they become synchronized between interlocutors within an interaction, also known as *Linguistic Style Matching (LSM),* which has been shown to predict a variety of group (Gonzales, Hancock, & Pennebaker, 2010; Sharma & Choundhury, 2018) and inter-personal (Ireland et al., 2011) outcomes.

Increased LSM might occur within communities for a variety of reasons, for instance, users abiding by the social norms of the group, (e.g. in profile practices such as bio length, actual photo for profiles), or linguistic features (e.g., use of slang or acronyms). This is particularly important for new users aiming to reduce dissimilarities between themselves and the group in order to integrate (Danescu-Niculescu-Mizil, West, Jurafsky, Leskovec, & Potts, 2013; Sharma & Choundhury, 2018). However, there are of course, exceptions to this trend, where some users might become innovators who shape and set new social norm trends. In contrast, other users may not accommodate their linguistic style at all, and instead maintain a idiosyncratic style despite the social pressure of the group (Danescu-Niculescu-Mizil et al., 2013). Hence, in the present research we examine whether users match (or not) their linguistic style to the norms within a specific online community sub-group. In doing so, we argue that if users are indeed matching their linguistic style to each of the sub-groups within a community, which in turn differ from each other, this could be seen as a reflection of social role shifting based on the extent to which they accommodate their linguistic style to match a group.

Additionally, there will also be differences in users' perception of social norms (e.g., what is acceptable behavior or communication in a group) that will further guide their behavior in groups. Across many online communities, and specifically, Reddit (which is examined here), it is common to have various types of social norms or 'rules' in order to moderate user behavior and to reduce harassment and unruly behavior (Matias, 2019). Rules and moderation can be important on reddit, where if posts violate moderation rules, it is immediately removed automatically or by a moderator with instructions regarding how to correct (the latter if it is a minor error such). However, should a user consistently break the rules or are seen as 'trolling' could lead to a user being banned from specific subreddits, reddit-wide bans, or a shadow ban (where users can up/downvote, but if they post or comment they are automatically deleted) (Reddit, 2019b). Therefore, understanding whether moderation actually impacts behavior and mitigates problematic or unruly behavior is important (Duggan, 2017; Jhaver, Ghoshal, Bruckman, & Gilbert, 2018). In the present study, we focus not only on whether users linguistic style shifts across different online communities, but also ask whether explicit differences in moderation has any impact on linguistic style accommodation. Hence, we seek to address the following research questions;

1. Do users accommodate their linguistic style to match the linguistic style of different subreddits?

2. Do explicit moderation ‘rules’ or social norms impact on user linguistic style matching with the community?

2 Methods

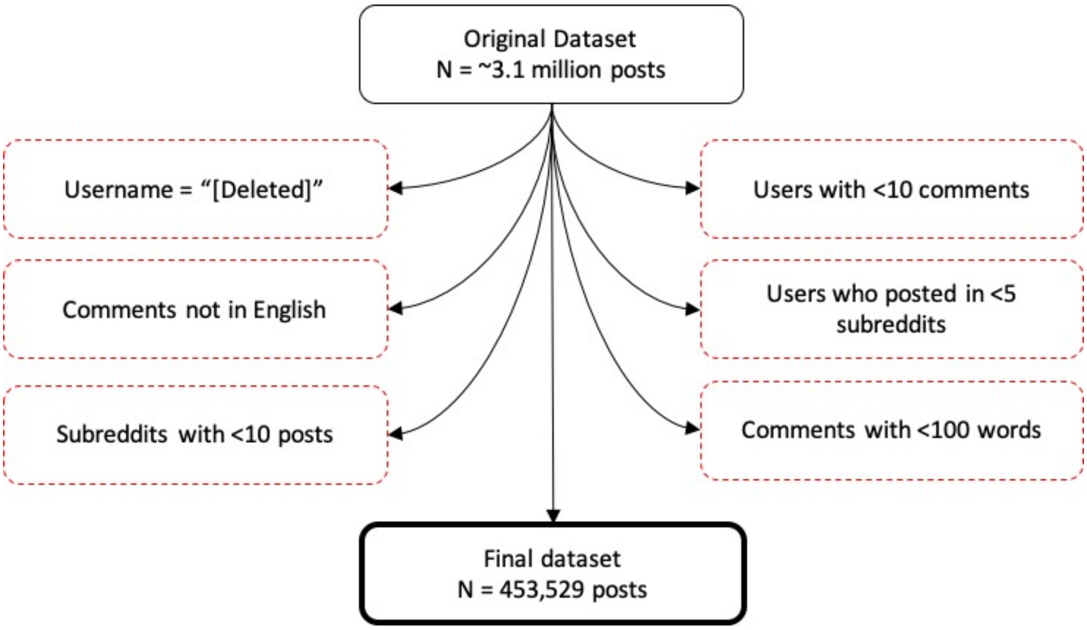
2.1 Dataset & Ethics

This study used secondary data that is publicly available (Syed, Voelske, Potthast, & Stein, 2018). From the original dataset, this paper used only the usernames, content data, and the subreddits in which the content was posted. The present study was reviewed and approved by the Ethical Implications of Research Activity (E1RA) process within the University of Bath, School of Management (reference number: 2320). IRB approval was not required for this work as it only utilized previously published secondary data sources.

2.2 Data Preprocessing

Results from early data analysis confirmed that the majority of users only post once, which is to be expected with the vast number of potential ‘throwaway’ accounts seen in online communities (e.g., De Choudhury & De, 2014; Leavitt, 2015; Pavalanathan & De Choudhury, 2015) (Appendices 1 and 2). As this study is interested in individual linguistic style, we made a series of decisions in order to reduce the dataset and to ensure there is sufficient data for each individual and each subreddit (Figure 1).

Figure 1 shows the six key decisions made to reduce the original dataset. This reduced the dataset from ~3.1 million posts down to N=453,529 posts, with N=24,180 individuals, and N=2,724 subreddits.



As shown in Figure 1, users who posted less than 10 times and in less than five subreddits were removed from the dataset. Similarly, subreddits in the data that were posted to less than 10 times were also removed. Posts with under 100 words were also removed, as well as any posts not in English. As we wanted to ensure there was sufficient data on each user to calculate a reliable measure of their linguistic style and LSM shift, we adopted a cut off of >1,000 words for each individual. Further, users who had removed their accounts or had their posts removed (denoted on Reddit as "[Deleted]" were also removed for this study. This reduced the dataset from ~3.1 million posts down to N=453,529 posts, with N=24,180 individuals, and N=2,724 subreddits.

To analyze language use, Pennebaker's LIWC (Pennebaker et al., 2015) was employed. LIWC is a 'bag of words' approach to computational linguistics that counts the proportion of words in a category (e.g. pronouns) relative to the total word count. The selected categories from LIWC can be found in Table 1, where we used the categories used for calculating Linguistic Style Matching (LSM) scores (Ireland et al., 2011). Table 1 also shows the overall rates of use (%) and standard deviation for all users (N=24,180).

Table 1. Usage Rates (%) and standard deviation (SD) for all users in dataset (N=24,180) for all LIWC categories included in analyses.

Word Category	Examples	Rate of Use (%)	SD
Function Words (overall)	–	54.00	5.42
Personal Pronouns	I, her, them	8.25	4.00
Impersonal Pronouns	It, that, anything	5.99	2.36
Articles	A, an, the	7.16	2.41
Prepositions	Under, off, in	13.43	2.54
Auxiliary Verbs	Will, shall, could	9.31	2.56
Conjunctions	Because, and, but	7.07	1.98
Negations	No, not, never	2.04	1.27
High Freq Adverbs	Very, rather, just	5.57	2.06
Quantifiers	Much, few, lots	2.78	1.49

3 Study One: Linguistic Style Matching Across Subreddits

3.1 Overview

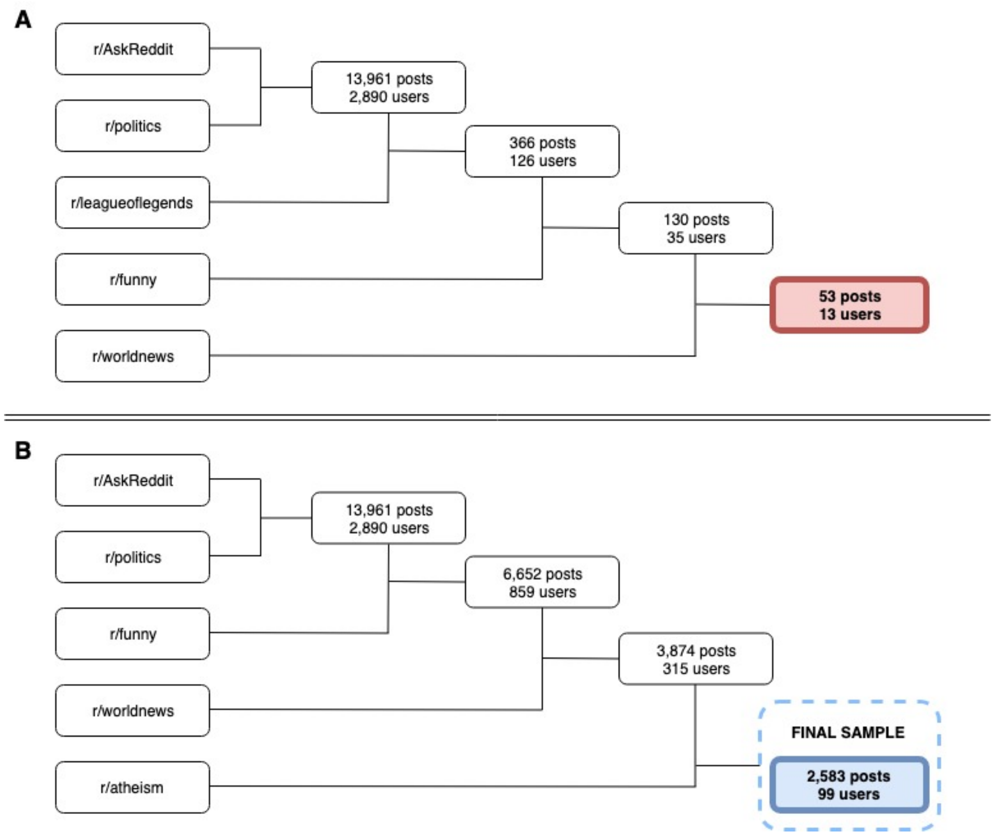
The aim of this study was to investigate users' linguistic behavior across different subreddits; therefore, we ran a within-subjects ANOVA in order to test if individuals significantly changed their linguistic style across subreddits. We used the nine dimensions of function words (Table 1) to calculate linguistic style for each user. We reduced the dataset to focus on a group of users who had all posted across the same subreddits in order to compare how users may or may not adapt their use of linguistic style to match each subreddit.

We first ordered the subreddits in terms of the highest number of posts in our dataset. From here, we selected five subreddits with users who have posted in all of them (details in Table 2). We removed highly-specific subreddits (e.g., r/leagueoflegends) as this greatly reduced the number of users and posts who had posted across these subreddits (Figure 1). In the final subset of data, there were 2,583 posts from 99 unique users who had all posted in r/AskReddit, r/politics, r/worldnews, r/funny, and r/atheism. These subreddits were also chosen as they cover a variety of topics from general advice, politics, to entertainment.

Table 2. Five subreddits used in Study One with community descriptions, number of posts, and number of unique users.

Subreddit	Description	# of Posts	# of Users
r/AskReddit	'To ask and answer questions that elicit thought-provoking discussions' (Reddit, 2019a)	60,925	15,199
r/politics	'for current and explicitly politics U.S. news' (Reddit, 2019b)	9,319	4,122
r/funny	Subreddit dedicated to humor, however, strictly no meme posting	6,043	4,009
r/worldnews	'About major news around the world, excluding US-internal news/ US politics' (Reddit, 2019d)	5,903	3,156
r/atheism	Subreddit for atheism and agnosticism discussion	5,658	2,701

Figure 2. Flow diagrams illustrating how participants were identified based on posts appearing across common subreddits. Diagram A consists of the five subreddits with the most posts in our dataset, however, only 13 users had posted in all five subreddits. Hence, we removed r/leagueoflegends due to its specificity. Diagram B shows the top five subreddits with the greatest number of posts in our dataset when excluding r/leagueoflegends, which has a final sample of 99 users. This sample is used in Study One.



3.2 'LSM Shift' Calculation

Linguistic style matching does not provide directionality in terms of shifting towards or away from another person or group, but rather provides an overall score (or 'match') between two observations of language use. We used the traditional method to calculate LSM (denoted below), then we developed an additional step to indicate 'LSM Shift' (LSM_S).

First, we considered each of the nine types of function words *separately* to calculate the three different LSM scores: LSM_A, LSM_B, and LSM_S. For each of these LSM scores, we used the typical calculation below, noting that the numerator uses *absolute values* as denoted by |x – x|). In the formulae below, personal pronouns (ppron) is used as the example.

$$\mathrm{LSM_{ppron}} = 1 - \frac{(|ppron_{user} - ppron_{subreddit}|)}{(ppron_{user} + ppron_{subreddit})}$$

LSM_A is a 'by chance' match, as this is naturally occurring baseline of concordance based on the user's average use of each type of function word from *all* their posts and the average use of each type of function word seen in each of the subreddits. Hence, each user (N=99) had five separate LSM_A scores between their average use of each function word and each of the five subreddit's average use of each type of function word.

LSM_B is the subreddit specific measure, where we filtered out any posts from users outside of the five subreddits of interest. Then, we calculated the average use of each type of function word for each user again based on their posts in the specific subreddits *only*. This was then matched to the subreddit average use of each type of function word (same as the subreddit score in the LSM_A calculation).

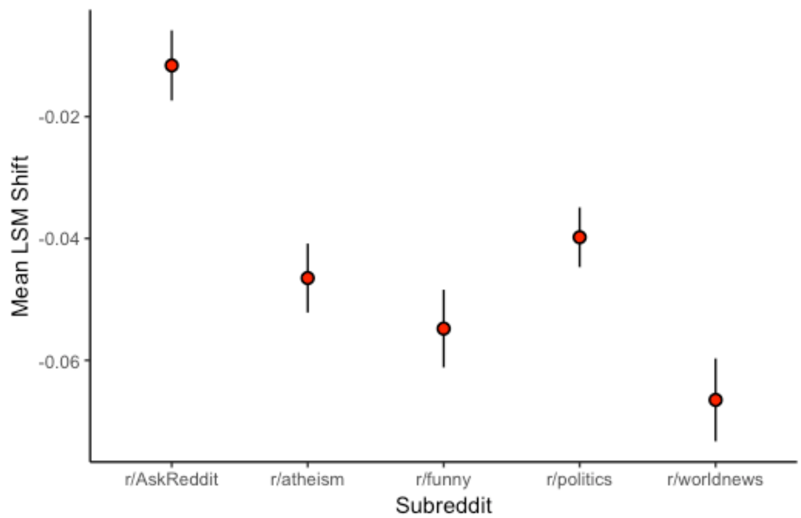
Finally, we compare the difference between the two:

$$\mathrm{LSM\_S} = \mathrm{LSM\_B} - \mathrm{LSM\_A}$$

At this point, we have nine versions of LSM_S – one for each of the nine function words. Typically, at this point, the average is taken in order to have an overall LSM score from the overall linguistic style.

We regard LSM_S (LSM Shift) as the increased or decreased matching taking into account naturally occurring matching between user's linguistic style and that of the subreddits. If it is positive number, this means the user shifts their linguistic style towards the subreddit and if it is negative, they are shifting away. Figure 3 shows the mean and standard error of LSM Shift in each subreddit. Additionally, Figure 4 is a visualization of the user LSM_S scores across the five subreddits.

Figure 3. Mean scores (red points) and standard error of Linguistic Style Matching Shift in each of the five selected subreddits (N=99). Y-axis denotes the average LSM Shift of users in each subreddit.



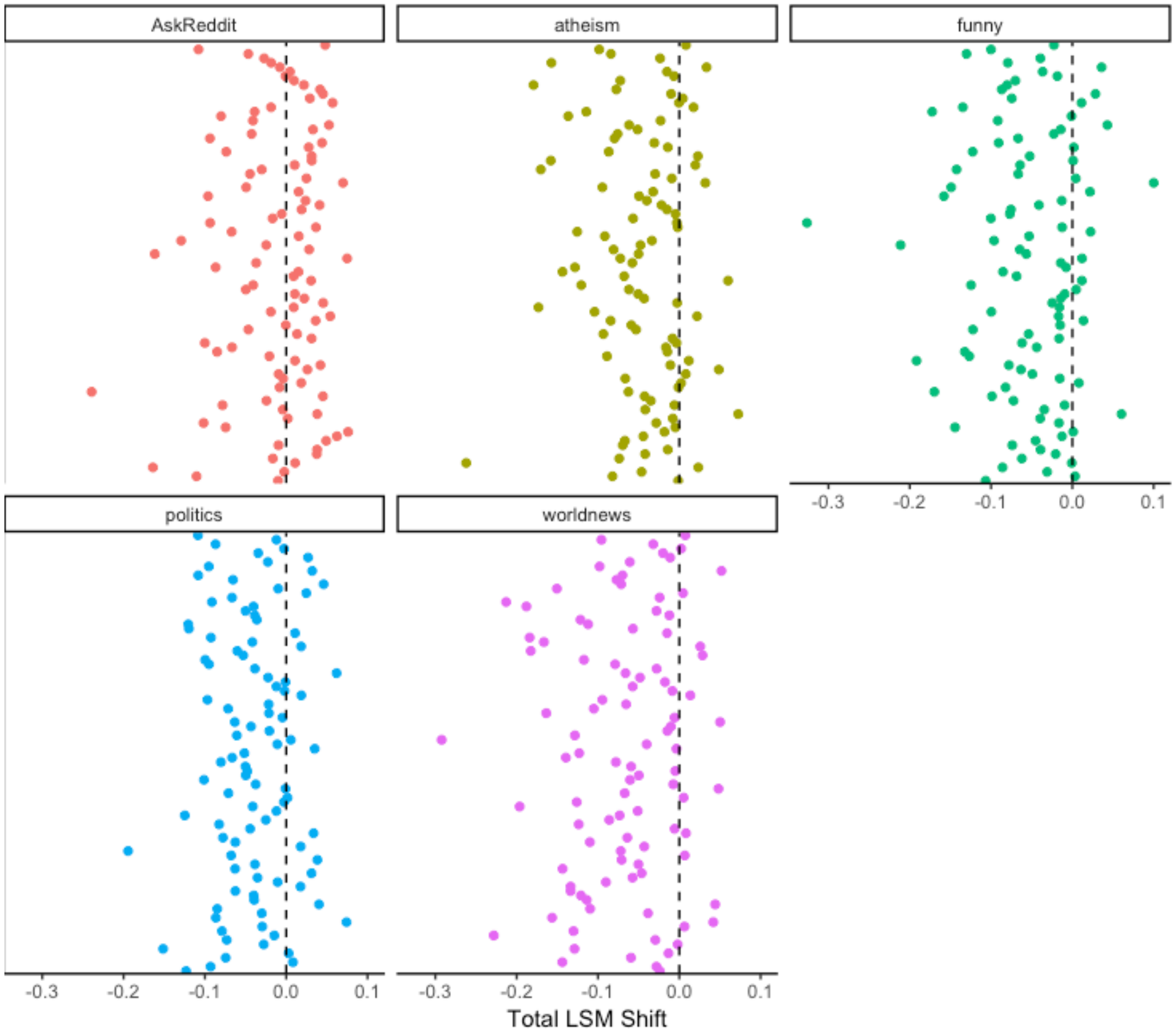
We note that on average, users shift *away* from the subreddit norms, which aligns with prior research findings that convergence is often characterized by the least amount of divergence when studying online communities (Jones et al., 2012). We ran a within-subjects ANOVA in order to test whether individuals significantly shifted in terms of their total LSM shift (LSM_S) in each of the five subreddits, $F(4, 392) = 15.4$, $p < .001$, $\eta p^2 = .136$. Essentially, this demonstrates there is a significant difference in the degree to which each user converges or diverges in LSM_S across different subreddits. We ran additional non-parametric (Friedman) tests as the data were not normally distributed. The Friedman's test replicated this result. In terms of these pairwise comparisons, we note moderate effect sizes (Table 3). The non-parametric test (Wilcoxon Rank) replicates the parametric findings, with the exception of funny/politics comparison was not significant p=.096.

Table 3. Pairwise comparisons (uncorrected) between LSM Shifts. Absolute Cohen's d (LHS) and Absolute Mean Differences (RHS) for Each Comparison.

	r/AskReddit	r/atheism	r/funny	r/politics	r/worldnews
r/AskReddit	-	.04	.04	.03	.06
r/atheism	$.47^{**}$	-	.01	.01	.02
r/funny	$.67^{**}$	.10	-	.02	.01
r/politics	$.42^{**}$	.10	$.22^{*}$	-	.03
r/worldnews	$.67^{**}$	$.25^{*}$	.15	$.37^{**}$	-

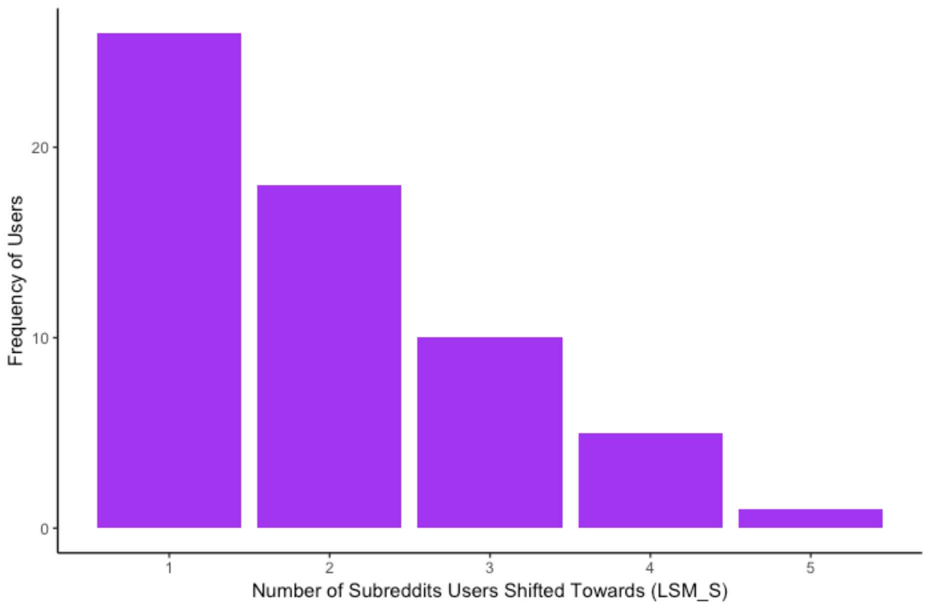
$^{*}p<.05$; $^{**}p<.00$

Figure 4. Each point denotes a participant; y-axis is each user. Participant LSM Shifts (LSM_S) based on their overall LSM Score (explained in the methods section). A positive number denotes a shift towards the group and a negative number denotes a shift away from the group). The dashed line denotes no movement (0).



From Figure 4, there are a number of users whose language *does* converge towards each of the subreddits. This was investigated further in order to see whether these were often the same 'chameleon'-like users. Figure 5 below shows that in fact, 60 out of the 99 users converged their linguistic style to match the subreddit's at least once. Only one user, however, converged their linguistic style towards all five of the subreddits, which we would regard as a true chameleon in this study. Typically, users converged towards r/AskReddit the most (41.88%), followed by r/politics (17.09%), r/funny (15.38%), and lastly, both r/worldnews and r/atheism (12.82%).

Figure 5. Frequency of unique users that shifted towards any of the five selected subreddits. In total 60 out of the 99 users shifted towards at least one subreddit (graphed below). 26 users moved towards only one subreddit and only 1 user shifted towards all five subreddits.



4 Discussion of Study One

We analyzed 99 users that posted in all of the five selected subreddits. We adapted the LSM calculation in order to calculate 'LSM Shift' for each user. This was based on the difference between their linguistic style in only the five subreddits of interest *and* their linguistic style based on all subreddits they had posted in (see above for LSM Shift Calculation for more details). This allowed us to see whether their linguistic style in specific subreddits converges or diverges from subreddits based on their typical linguistic behavior.

We found that on average, users tended to shift away from all subreddits, which aligns with the findings of Jones et al. (2012) that divergence is normative in online communities (most likely due to the scale and diversity of content generated in online groups). However, we note that out of those five subreddits, users often accommodated the most in r/AskReddit and the least for r/atheism and r/worldnews. We ran a repeated-measures ANOVA in order to examine whether individuals significantly shift their linguistic style across different subreddits, which was indeed significant. This demonstrates that often, linguistic style does indeed change based on the context. Additionally, the effect sizes found were moderate in size, and are comparable to other LSM research (e.g., Cox & Kersten, 2016; Gonzales et al., 2010; Ireland et al., 2011).

We found that many users converged their linguistic style to match the community, and similarly many others diverged (Figure 4). We investigated this further in order to understand whether this was perhaps just noise, or whether there were consistent users who did adapt their linguistic style. As demonstrated in Figure 5, we found that there were many users who displayed some form of 'shape shifting' or chameleon-like behavior (Jones et al., 2012), where they often converged their linguistic style to match the community's. Typically, we found this occurred in r/AskReddit. However, we note there were many users who did not display this 'shape shifting' behavior and therefore did not converge across all or many subreddits, which aligns with prior research and is not uncommon (Jones et al., 2012).

This variance in LSM 'Shift' suggests there is a mixture of individual differences and subreddit (context) norms impacting on individuals. For instance, users tended to accommodate in r/AskReddit far more than r/atheism or r/worldnews, which may reflect how users change and adapt their behavior according to the environment around them (e.g., Fiske, 2010). We can understand the variance in behavior via CAT, where people will converge or diverge their communication according to a variety of factors from identification with the group, desire to integrate, to impression management. For instance, r/AskReddit is open for discussion as it allows for comments and posts across all topics from entertaining questions through to serious questions about work place or relationship difficulties – where users might seek to bond with others (Ren et al., 2007) demonstrated by a convergence of linguistic style. On the other hand, r/politics and r/worldnews only allow for discussion on specific topics, which are likely to contain conflicting opinions (Major, Kaiser, O'Brien, & McCoy, 2007) and so people accommodate less as a result. Sharma and De Choudhury (2018), also observed that LSM can be a reflection of support within communities, where user 'roles' within r/AskReddit might indeed be more supportive than when they are engaging with political discussion, in comparison. This highlights the potential differences in use of online communities, where those in r/AskReddit might be looking for a community to bond with and to have support from (e.g., common bond) whereas subreddits like r/politics, users will likely have a strong political identity they hold, which may impact on how they interact with others who have the same or opposing political identity (Ren et al., 2007).

Without engaging and analyzing the comments more closely, specific examples of and reasons why users accommodate less in certain subreddits is difficult to pinpoint here – especially as r/AskReddit is indeed a general community rather than specifically formed for advice. These findings do however, align with prior literature demonstrating that users do change their roles in online communities (Davidson et al., 2019), which is also reflected by adaptation of linguistic style (Danescu-Niculescu-Mizil et al., 2013; Muir, 2016) to either converge or diverge from the community (Giles & Ogay, 2007).

In a second analysis, we test whether explicit differences in moderation impacts on user linguistic style matching to subreddits of a similar nature. This is to ensure that these differences in linguistic style demonstrated in Study One is not simply noise or is driven by the vastly different topics discussed across r/AskReddit, r/politics, r/funny, r/worldnews, and r/atheism. This aims to shed further light on how users accommodate linguistic style across online contexts, which is important for marketing and even security purposes. For instance, if user behavior is inconsistent in similar contexts, any marketing and advertising messages – or indeed behavioral nudges, should therefore be adapted to suit that platform in order to maximize any potential impact. This also demonstrates how complicated user behavior is online (as with offline), which provides an abundance of opportunities for future research, as this work merely utilizes content data.

5 Study Two: How Does Community Moderation Impact LSM Shift?

We now examine whether the variance in LSM shift observed in Study One could be explained by the variety of content and topics discussed in the five selected subreddits (r/AskReddit, r/atheism, r/funny, r/politics, and r/worldnews). It has been demonstrated that explicitly stated social norms in online communities are important for understanding and predicting user behavior (Matias, 2019), therefore, comparing similar subreddits that are explicitly moderated differently is a way to test whether this does indeed have an impact on user behavior based on a user's typical linguistic style. Matias (2019) demonstrated that posting rules on reddit increased the likelihood of joining conversationalists posting, and increases the chances of their new posts to not be removed by moderators for breaking rules, as the social norms are more explicitly enforced and shown. Abiding by rules is important on reddit, as if a post breaks the moderation rules, it is immediately removed automatically or by a moderator with instructions regarding how to correct it if it was a minor error (e.g., leaving out age and gender on a post). Breaking the rules consistently could lead to a user being banned from the subreddit or eventually locking the user's account. Understanding whether moderation tactics actually impact on behavior is critically important for online communities in order to address online harassment and unruly behavior as this continues to be problematic (Duggan, 2017; Jhaver, Ghoshal, Bruckman, & Gilbert, 2018) for a variety of communities from Reddit to Facebook, Twitter, and beyond.

5.1 Participants

In this study, we selected a new sample of participants. This is because we seek to analyze whether user linguistic behavior changes across subreddits of a *similar* nature (e.g., subreddits with the same topic and type of content), unlike Study One. Additionally, we needed to find subreddits with clear differences in moderation (e.g., a subreddit with explicit rules and moderation warnings versus a subreddit with little to no rules and moderation of content).

From *Study One*, r/politics explicitly has strict rules of moderation, where users are expected to '*be civil*' (Reddit, 2019c). In order to examine whether moderation may impact on a user's LSM Shift as demonstrated in *Study* One, we searched for another political subreddit with little to no moderation of content. Hence, we used r/worldpolitics as this is reddit's '*free speech political subreddit*' and effectively allows most content including: '*offensive content, fake news, propaganda,'* (Reddit, 2019f). In contrast, we also analyzed two sets of subreddits that are similar in content, but do *not* have obvious differences in moderation and rules. Alongside the political subreddits (r/politics and r/worldpolitics), we examined two large gaming (r/gaming and r/pcgaming) and relationship advice (r/relationships and r/relationship_advice) subreddits. The gaming theme was chosen as r/gaming is the third largest (safe for work ('SFW')) subreddit (Reddit, 2019d) and is a common topic or hobby seen online. However, we wanted to consider general gaming subreddits, for example, r/gaming, rather than highly specific gaming subreddits (e.g., r/leagueoflegends). Similarly, the relationships theme was chosen as this is another common topic on reddit, as reflected by r/relationships being a top ranked subreddits (#23 out of over 1.2 million subreddits (Richter, 2017) for high recent activity (Reddit, 2019d) upon time of writing.

Similar to *Study One*, we ensured a reasonable number of posts and unique users had posted in both subreddits for each theme. When sourcing subreddits of similar topics that users had posted across, the number of posts and unique users tended have insufficient numbers when looking across three similar subreddits. For example, the gaming subreddits (r/gaming and r/pcgaming) together have 101 unique users, however, when r/gamingpc was also included this fell to one unique user. Similarly, r/politics and r/worldpolitics had 72 unique users posting in both, however, the addition of r/neutralpolitics caused this number of unique users to fall to three. Hence, we proceeded with sets of two subreddits to compare. Table 4 shows the three new datasets for Study Two.

Table 4. Details of Datasets of Subreddits for Study Two. Each theme consists of two subreddits, which contains users who posted in both. Hence, in the political subreddits, there were 72 users who posted in both r/politics and r/worldpolitics. In these subreddits, there were 239 posts.

Theme	# of Posts	# of Users
Political (moderated vs not moderated) *r/politics & r/worldpolitics*	239	72
Relationships (no moderation differences) *r/relationships and r/relationship_advice*	1,411	272
Gaming (no moderation differences) *r/gaming and r/pcgaming*	214	101

5.2 Results

For each of the users in each of the datasets (Table 4), we calculated their LSM_S score (detailed in Study One). This LSM_S score was then used again to compare users across each of the subreddits. For each of the datasets, we are comparing user LSM Shift across two subreddits, therefore we employed t-tests.

First, we checked for violations of normality, which only the relationships sample violated (Shapiro-Wilks p=.025), therefore we ran non-parametric tests (Wilcoxon Rank) alongside the T-test for the r/relationships and r/relationship_advice T-Test.

Both the relationships and gaming subreddits were *not* significantly different from one another [all p's >.15]. However, LSM shift *was significant* across r/politics and r/worldpolitics, [$t(71) = 3.68$, $p<.001$, $d = .434$], mean difference = .033 (Figures 6 and 7). Mean difference is reported for completeness and demonstrates the absolute difference between each group mean (r/politics and r/worldpolitics).

Figure 6. 6A shows the mean and median values from r/politics and r/worldpolitics to demonstrate the clear difference in LSM Shift for users. 6B shows the scatter plot showing each individual user's post in both r/politics and r/worldpolitics (y axis is each individual user) in terms of LSM Shift. A positive number denotes a shift towards the group and a negative number denotes a shift away from the group). The dashed line denotes no movement (0).

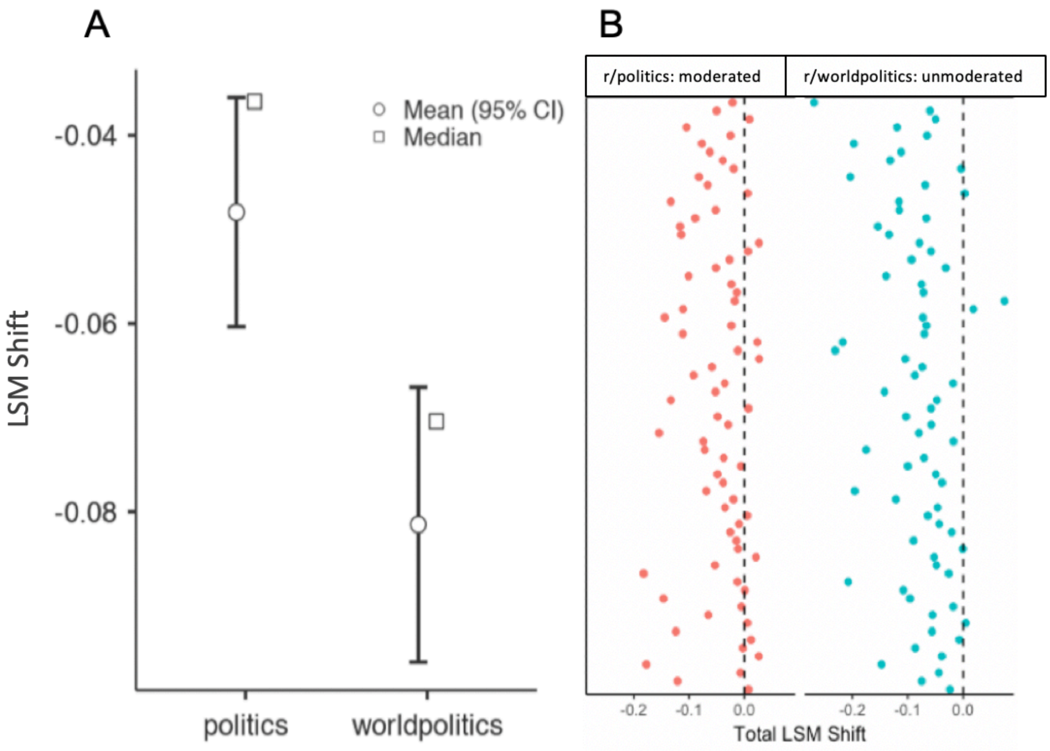
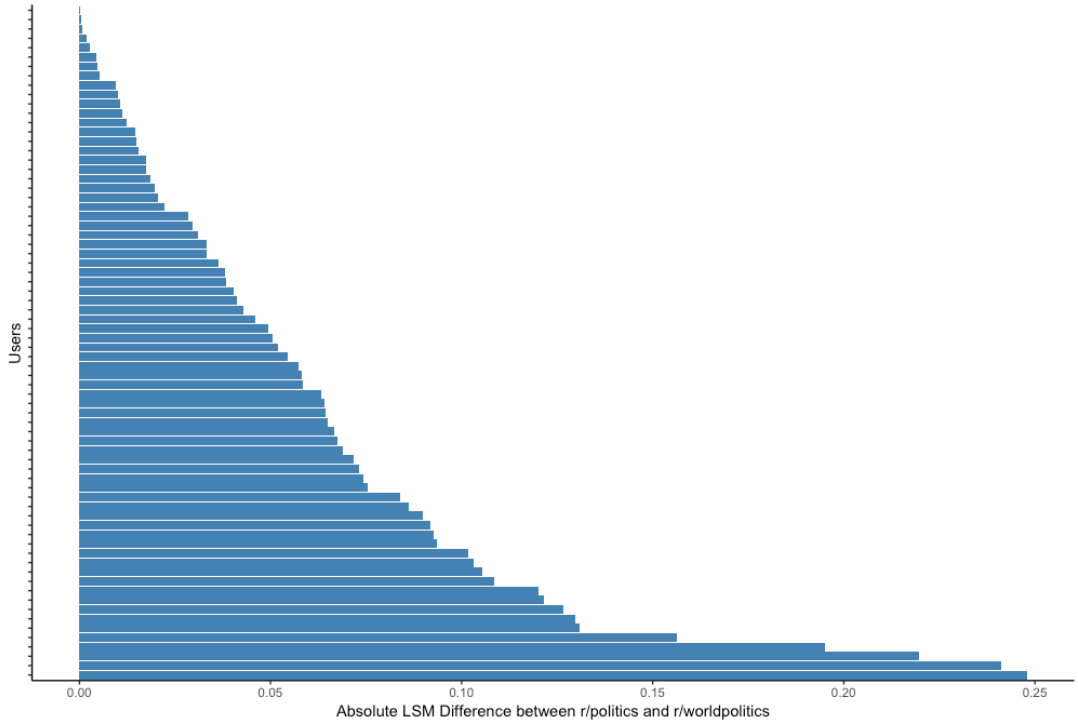


Figure 7. Absolute difference in LSM Shift score for each user in r/politics and r/worldpolitics (N = 72). Absolute difference was used as this graph aims to show how many users shift their linguistic style in different subreddits. Larger absolute difference LSM Scores denote larger differences in linguistic style across r/politics and r/worldpolitics – or 'chameleon' like behavior. Note: y-axis denotes each user; usernames are hidden to maintain anonymity.



6 General Discussion

Across two studies, we analyzed multiple subsets of redditors who had posted across the multiple communities. *Study One* captured a broader spectrum of subreddits, where we analyzed r/AskReddit, r/atheism, r/funny, r/politics, and r/worldnews. After observing that user LSM Shift was significantly different across subreddits in *Study One* and finding that there were many users who did indeed display ‘shape shifting’ behaviors, where they converged their linguistic style to match the subreddit’s multiple times. This is an active demonstration of CAT, where users change and adapt their communication style both towards and away from communities. *Study Two* intended to consider whether users adapt their linguistic style and shift across similar subreddits. This is because we wanted to distinguish between whether these differences in *Study One* may be partially explained by the topics and purpose of these subreddits changing substantially. For instance, as users change significantly between r/AskReddit and r/worldnews, how much of that difference relates to the community norms itself?

Study Two analyzed a further three sets of two similar subreddits in terms of topic: political (r/politics and r/worldpolitics), relationship advice (r/relationships and r/relationship_advice), and gaming (r/gaming and r/pcgaming). Therefore, in *Study Two* we collected three groups of two similar subreddits to compare user linguistic style. Additionally, we wanted to consider whether explicit differences in moderation may impact on user linguistic style. We found that users *did* shift their linguistic style significantly across political subreddits, however, this was not the case for relationship advice or gaming subreddits. This could relate to the explicit differences in moderation on these subreddits, which aligns with the work of Matias (2019) demonstrating posting community moderation rules onto the discussion pages positively impacted newcomer’s user behavior.

Here we compared r/politics, which aims to be civil and to have fair and balanced discussions, to r/worldpolitics that allows any content that includes ‘*offensive content, fake news, propaganda, feature stories*’ etc., as long as it does not violate the overall rules of reddit (Reddit, 2019f). Our results showed that user LSM ‘Shift’ was significantly less in r/worldpolitics in comparison to r/politics. Comparing our findings to Matias (2019), this was not unexpected with the explicit lack of moderation on r/worlpolitics. We can consider this in terms of CAT theory, where perhaps the lack of needing to manage their impression due to the lack of moderation can explain the larger divergence in LSM 'Shift'. Further, we suggest that r/worldpolitics fosters a community with a reduced necessity of self-regulatory behavior, which additionally helps to explain the high levels of communication divergence (Joinson, 2001; Joinson & Paine, 2007). Hence, this aligns with perspectives from Joinson (2001), who proposed factors, such as anonymity and lack of moderation would impact behavior. For instance, users might be more willing to express viewpoints they wouldn't otherwise in offline or less anonymous places, which may be demonstrated by many users shifting their LSM away from subreddit norms seen in both *Study One* and *Two*.

In addition, these findings have relevance to Lampe et al.'s (2014) work focusing on the importance of distributed moderation in online forums. This particularly necessary in political discussion forums, as there are increasing concerns regarding mis/disinformation. However, moderation could provide a level of quality assurance for the comment rather than solely focusing on the political position of a political comment (Lampe et al., 2014). As demonstrated by Matias (2019) and in the present study, explicit and strict moderation does have some impact on user behavior. However, there is much work to be done in this field building on the work of other scholars interested in online moderation (Matias, 2019; Wise, Hamman, & Thorson, 2006). This could be considered from both a social role or social identity perspective, or a more general lens via CAT.

Additionally, there is substantial variation in users' linguistic style, which is demonstrated by user shifts in communication behaviors across subreddits more generally across both studies. This is particularly apparent in Figure 7, where we can see the difference in LSM Shift for each user in r/politics and r/worldpolitics. In this figure, we can see that many users' LSM Shift with both communities was extremely different, while other users' amount of convergence or divergence was virtually indistinguishable (e.g., points closer to 0) across r/politics and r/worldpolitics. This is an important finding as there is further work to be done to understand how and why users change their communication behaviors, as well as to understand why some users are more flexible in their linguistic style than users.

We have shown that largely, users are different across subreddits, however, when comparing behavior across similar subreddits, user adaptation was not significant consistently across these communities. This therefore shows that while users will adapt their linguistic style, there is indeed a limit. Hence, how much they will adapt to the social norms and context is critically important to predict how a user may behave. This may have implications for security settings for the identification of users across different online contexts. For instance, having an awareness that linguistic style can, and does, change online – therefore, utilizing other types of data and analysis will be required to match users across settings. This behavioral flexibility will have marketing, advertising, and business implications (Miles, 2014). For example, we suggest that the way in which businesses advertise should be adapted according to the context and the potential or expected roles of their customer-base.

Finally, our findings demonstrate how *diverse* and *unique* an experience each user creates for themselves within Reddit. Similarly, despite there being over 1.2 million subreddits (Richter, 2017) and undoubtedly many relating to very similar topics as seen in in *Study Two*, users potentially find out subreddit and stick to it rather than subscribing to many of the same topic. This will likely translate across all forms of social media, where one user's experience is drastically different and indeed unique to them in comparison to other even similar users (e.g., users with similar interests, hobbies) (Belk, 2013).

6.1 Limitations

First, we only considered a single online platform, however, this was large and each individual subreddit is its own online community, therefore differences across subreddits may indeed be smaller than if we examined user linguistic style matching to a subreddit and the same user on a board on 4Chan, or any other online forum. Further, the use of function words (linguistic style and LSM) rather than content-based words helps mitigate issues with context, where we would expect to see similar variance in user behavior across different communities demonstrated here. Second, we acknowledge the number of users examined here is small (N = 72 – 272) in comparison to the overall dataset (N = 24,180). However, these sample sizes are comparable or larger than other LSM studies (e.g., Gonzales et al., 2010; Ireland et al., 2011; Muir et al., 2016; Niederhoffer & Pennebaker, 2002). There is also some debate as to whether LSM is sufficient for understanding user behavior shifts, as typically, the scores for LSM before calculating LSM Shift was high, and therefore, the chances for users to converge further, was often small. Hence, it was more likely for them to diverge by chance. This therefore demonstrates that while this provides an exploratory insight into how user linguistic style does shift, in order to solidify these claims, more in-depth analysis of content and potentially the use of other data types (e.g., metadata) will be required. Of course, as LSM was able to demonstrate these user shifts in behavior, we might anticipate that other methods would reveal even larger shifts in behavior and communication. Finally, the use of LSM remains a high-level approach, as it takes the average of user's use of all nine types of function words. There could be more subtle variation and fluctuations occurring linguistically, which warrants further studies. While it is clear users can, and do, change their use of language across subreddits, it is not yet clear whether there are more subtle changes in language are occurring, which requires further research, as above, other more intricate linguistics methods are likely to show larger differences across contexts. However, the key advantage of using LSM is that the outputs and analysis are generally straightforward to communicate and provides a solid foundation for considering later, more subtle variations in language.

7 Conclusion

While it is well-documented that people change and adapt their behavior across various contexts, little research has sought to examine this across online communities. This, to the best of our knowledge, is the first study to examine users across multiple communities to demonstrate linguistic accommodation to match (or not) with various communities. We found that users often accommodate their communication style according to the context, however, this was inconsistent. Akin to prior research, many users did not accommodate their communication style (Jones et al., 2012). However, we note there were several users that displayed 'chameleon' like behavior, where their communication style matched with some communities, which could be seen as a reflection of a shift in social role. However, there is a limit to this social flexibility, where the same users posting in subreddits of a similar nature were not always significantly different. Our results additionally demonstrate that moderation plays a role in the extent to which a user accommodates their communication style in communities. In all, this paper revealed how dynamic behavior is online across multiple contexts. Hence, moving forward, research that considers people across multiple contexts is becoming increasingly important when it comes to understanding individuals, groups and their social roles in a digitized society.

References

Ang, C. S., & Zaphiris, P. (2010). Social Roles of Players in Mmorpg Guilds. *Information, Communication & Society*, *13*(4), 592–614. https://doi.org/10.1080/13691180903266952

Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). The Impact of Community Commitment on Participation in Online Communities. *Information Systems Research*, *22*(4), 841–854. https://doi.org/10.1287/isre.1090.0265

Belk, R. W. (2013). Extended Self in a Digital World. *Journal of Consumer Research*, *40*(3), 477–500. https://doi.org/10.1086/671052

Chan, J., Hayes, C., & Daly, E. M. (2010). Decomposing Discussion Forums using User Roles. *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*, *112*(988), 215–218. Retrieved from http://aran.library.nuigalway.ie/xmlui/handle/10379/1114

Cox, C. R., & Kersten, M. (2016). Mortality salience increases language style matching and well-being. *Self and Identity*, *15*(4), 452–467. https://doi.org/10.1080/15298868.2016.1154888

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013). No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities. In *In Proceedings of the 22nd international conference on World Wide Web* (pp. 307–318). ACM.

Davidson, B. I., Jones, S. L., Joinson, A. N., & Hinds, J. (2019). The evolution of online ideological communities. *PLOS ONE*, *14*(5), e0216932. https://doi.org/10.1371/journal.pone.0216932

De Choudhury, M., & De, S. (2014). Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *In Eighth International AAAI Conference on Weblogs and Social Media.* (pp. 71–80).

Duggan, M. (2017). Online Harassment 2017. Retrieved May 6, 2019, from https://www.pewinternet.org/2017/07/11/online-harassment-2017/

Ellis, D. A. & Merdian, H. L. (2015). Thinking Outside the Box: Developing Dynamic Data Visualizations for Psychology with Shiny. *Frontiers in Psychology. 6:1782*. DOI: 10.3389/fpsyg.2015.01782

Fiske, S. T. (2010). *Social Beings* (2nd ed.). NJ, USA: Wiley.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context and consequence. In *Contexts of accommodation* (pp. 1–68). Cambridge, UK: Cambridge University Press.

Giles, H., & Ogay, T. (2007). Communication Accommodation Theory. In *Explaining communication: Contemporay theories and exemplars* (pp. 293–310). Mahwah, NJ: Lawrence Erlbaum.

Gleave, E., Welser, H. T., Lento, T. M., & Smith, M. A. (2009). A Conceptual and Operational Definition of ' Social Role ' in Online Community. *Proceedings of the 42nd Hawaii International Conference on System Sciences*, 1–11. https://doi.org/10.1109/HICSS.2009.6

Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language Style Matching as a Predictor of Social Dynamics in Small Groups. *Communication Research*, *37*(1), 3–19. https://doi.org/10.1177/0093650209351468

Heer, J. & Kandel, S. (2012). Interactive Analysis of Big Data, *XROS*, 19(1), 50-54

Hogg, M., Terry, D., & White, K. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*. https://doi.org/10.2307/2787127

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language Style Matching Predicts Relationship Initiation and Stability. *Psychological Science*, *22*(1), 39–44. https://doi.org/10.1177/0956797610392928

Jhaver, S., Ghoshal, S., Bruckman, A., & Gilbert, E. (2018). Online harassment and content moderation: The case of blocklists. *Transactions on Computer-Human Interaction (TOCHI)*, *25*(2). https://doi.org/10.1145/3185593

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, *31*, 177–192.

Joinson, A. N., & Paine, C. B. (2007). Self-disclosure, privacy and the Internet. In *Oxford handbook of Internet psychology* (pp. 237–252).

Jones, S., Cotterill, R., Dewdney, N., Muir, K., & Joinson, A. (2012). Finding Zelig in Text: A Measure for Normalising Linguistic Accommodation. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers* (pp. 455–465).

Lampe, C., Zube, P., Lee, J., Park, C. H., & Johnston, E. (2014). Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly*, *31*, 317–326.

Leavitt, A. (2015). This is a throwaway account: Temporary technical identities and perceptions of anonymity in a massive online community. In *Proceedings of the*

18th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 317–327). ACM.

Major, B., Kaiser, C. R., O'Brien, L. T., & McCoy, S. K. (2007). Perceived Discrimination as Worldview Threat or Worldview Confirmation: Implications for Self-Esteem. *Journal of Personality and Social Psychology*, *92*(6), 1068–1086. https://doi.org/10.1037/0022-3514.92.6.1068

Matias, J. N. (2019). Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 201813486. https://doi.org/10.1073/pnas.1813486116

Miles, D. A. (2014). Measuring customer behavior and profitability: Using marketing analytics to examine customer and marketing behavioral patterns in business ventures. *Academy of Marketing Studies Journal*, *18*(1), 141–165.

Muir, K. (2016). Power and linguistic accommodation in cmc. In *In the proceedings of the 66th International Communication Association Annual Conference*. Fukuoka, Japan.

Muir, K., Joinson, A., Cotterill, R., & Dewdney, N. (2016). Characterizing the Linguistic Chameleon: Personal and Social Correlates of Linguistic Style Accommodation. *Human Communication Research*, *42*(3), 462–484. https://doi.org/10.1111/hcre.12083

Muir, K., Joinson, A., Cotterill, R., & Dewdney, N. (2017). Linguistic Style Accommodation Shapes Impression Formation and Rapport in Computer-Mediated Communication. *Journal of Language and Social Psychology*, *36*(5), 525–548. https://doi.org/10.1177/0261927X17701327

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, *21*(4), 337-360+454. https://doi.org/10.1177/026192702237953

Papacharissi, Z. (2002). Presentation of Self In Virtual Life: Characteristics of Personal Home Pages. *Journalism & Mass Communication Quarterly*, 643–660.

Pavalanathan, U., & De Choudhury, M. (2015). Identity management and mental health discourse in social media. In *Proceedings of the 24th International Conference on World Wide Web*. (pp. 315–321). ACM.

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic inquiry word count: LIWC2015*. *Pennebaker Conglomerates*. Austin, Texas, USA. Retrieved from www.LIWC.net

Pennebaker, J. W. (2011). *The secret life of pronouns*. New Scientist, 211(2828), 42-45

Pfeil, U., Svangstu, K., Ang, C. S., & Zaphiris, P. (2011). Social Roles in an Online Support Community for Older People. *International Journal of Human-Computer Interaction*, *27*(4), 323–347. https://doi.org/10.1080/10447318.2011.540490

Preece, J., & Schneiderman, B. (2009). The Reader-to-Leader Framework: Motiviating Technology-Mediated Social Participation. *Transactions on Human-Computer Interaction*, *1*(1), 1–25. https://doi.org/10.5121/ijfcst.2014.4403

Ransbotham, S., & Kane, G. K. (2011). Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia. *MIS Quarterly*, *35*(3), 613–627.

Reddit. (2019a). AskReddit Wiki. Retrieved March 18, 2019, from https://www.reddit.com/r/AskReddit/wiki/index

Reddit. (2019b). Frequently Asked Questions. Your reddit account. I've been banned. Why? How do I get unbanned? Retrieved June 22, 2019, from https://www.reddit.com/r/help/wiki/faq

Reddit. (2019c). Politics Wiki. Retrieved March 18, 2019, from https://www.reddit.com/r/politics/wiki/index

Reddit. (2019d). RedditList. Retrieved April 9, 2019, from http://redditlist.com/

Reddit. (2019e). World News Rules and Guidelines. Retrieved March 18, 2019, from https://www.reddit.com/r/worldnews/wiki/rules

Reddit. (2019f). World Politics. Retrieved April 5, 2019, from https://www.reddit.com/r/worldpolitics/

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds. *MIS Quarterly*, *36*(3), 841–864.

Ren, Yuqing, Kraut, R., & Kiesler, S. (2007). Applying common identity and bond theory to design of online communities. *Organization Studies*, *28*(3), 377–408. https://doi.org/10.1177/0170840607076007

Richter, F. (2017). The Explosive Growth of Reddit's Community. Retrieved April 9, 2019, from https://www.statista.com/chart/11882/number-of-subreddits-on-reddit/

Schmader, T., & Sedikides, C. (2018). State Authenticity as Fit to Environment: The Implications of Social Identity for Fit, Authenticity, and Self-Segregation. *Personality and Social Psychology Review*, *22*(3), 228–259. https://doi.org/10.1177/1088868317734080

Sharma, E., & Choundhury, M. De. (2018). Mental Health Support and its Relationship to Linguistic Accommodation in Online Communities. In *CHI '18 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Montreal, QC, Canada. https://doi.org/10.1145/3173574.3174215

Steele, J. & Iliinsky, N. (2010). Beautiful visualization: looking at data through the eyes of experts. O'Reily Media, Inc. Sebastopol, Canada.

Syed, S., Voelske, M., Potthast, M., & Stein, B. (2018). Dataset for generating TL;DR [Dataset]. https://doi.org/10.5281/zenodo.1168855

Turner, T., Smith, M., Fisher, D., & Welser, H. (2005). Picturing Usenet: Mapping Computer-Mediated Collective Action. *Journal of Computer-Mediated Communication*, *10*(4), 00.

Valero-Mora, P. & Ledesma, R. (2014). Dynamic-Interactive Graphics for Statistics (26 Years Later), *Current Topics in Statistical Graphics, 37(2)*, 247-260

Welser, H., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. *Proceedings of the 2011 IConference on - IConference '11*, 122–129. https://doi.org/10.1145/1940761.1940778

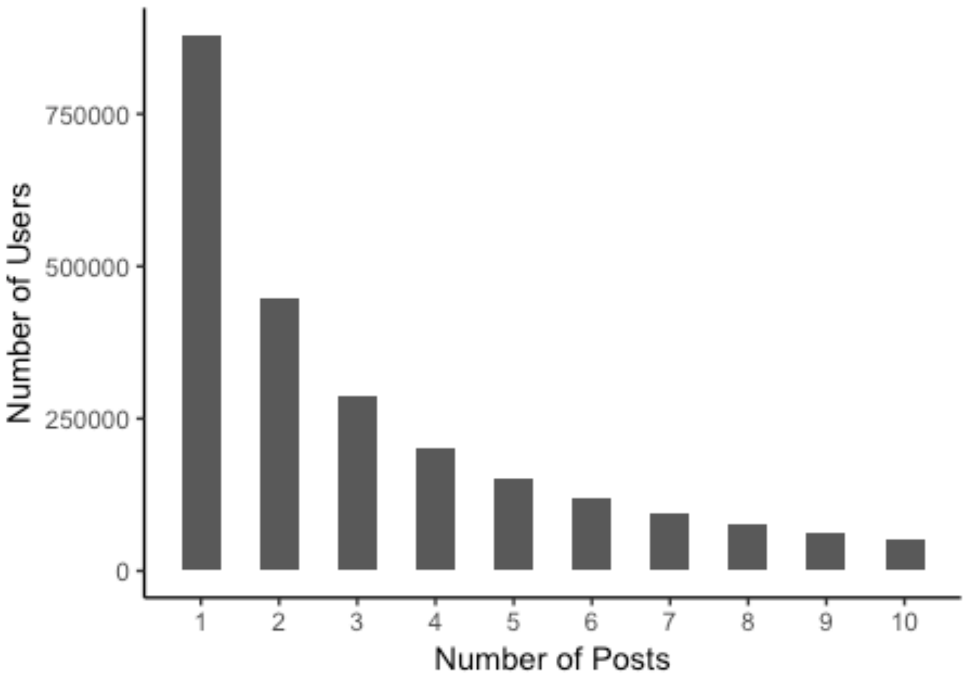
Welser, H. T., Gleave, E., Fisher, D., & Smith, M. A. (2007). Visualizing the Signatures of Social Roles in Online Discussion Groups. *Journal of Social Structure*, *8*(2), 1–32.

Wise, K., Hamman, B., & Thorson, K. (2006). Moderation, response rate, and message interactivity: Features of online communities and their effects on intent to participate. *Journal of Computer-Mediated Communication*, *12*(1), 24–41.

Appendix

Appendix 1.

Histogram of Number of Posts from Pre-Processed dataset. Histogram of those who posted between 1 and 10 times



Appendix 2. Consistency and Inconsistency of Users and Subreddit's Use of Function Words (overall)

Individual User Linguistic Style (In)consistency

For the final part of the initial data exploration and visualization, we wanted to understand the individual variance of function words for each redditor (Table A). This provides an initial indication of flexible users are, and whether they will accommodate their linguistic style across various subreddits. Here, we used z-scores in order to calculate this.

Table A. Z-scores of function words for individual redditors (N = 24,180). We focus only on the frequency of users who fall >1.96 sd from the mean (95% confidence interval) and 3sd from the mean to show more 'extreme' inconsistent users based on their variance of use of each type of function word. We only look at inconsistence (e.g., >1.96 or 3 sd from the mean, rather than extremely consistent users <1.96 or 3).

Word Type	>1.96 sd from *M*	% of Users	>3 sd from *M*	% of Users
Function Words (Overall)	876	.036	403	.017
Personal Pronouns	1,122	.046	213	.009
Impersonal Pronouns	1,024	.042	328	.014
Articles	1,057	.044	330	.014
Prepositions	973	.040	262	.011
Auxiliary Verbs	1,085	.045	286	.012
Conjunctions	607	.025	145	.006
Negations	917	.038	316	.013
High Frequency Adverbs	785	.012	202	.008
Quantifiers	1,041	.043	323	.013
Number of Individual Users	24,180			

The z-scores show that individuals are extremely consistent with their use of function words, as so few redditors fall further than >1.96 standard deviations from the average variance of all redditors in terms of function word use. This aligns with prior work demonstrating that the use of function words is indeed consistent (e.g., Boyd & Pennebaker, 2017; Niederhoffer & Pennebaker, 2002). However, this finding does raise the question as to whether redditors in general will indeed accommodate their linguistic style at all across different subreddits.

In response, we conducted some further preliminary data visualization, where we looked at 200 users: 100 users with the highest variance of function word use (Figure A) and the 100 users with the lowest variance of function word use (Figure B) in the dataset. While this is only a small percentage of the overall number of users, it does however, provide some insight into how varied user's function word (overall) use might be.

*Figure A. Top 100 users with the **highest** variance of overall function word use (%). Each purple dot and line denotes a single unique user. Variance was calculated based on all user's posts within the dataset (post-cleaning). The purple dot denotes the user's average number of function words across all of their posts. The error bar denotes the highest and lowest % function words across all posts. x-axis shows users.*

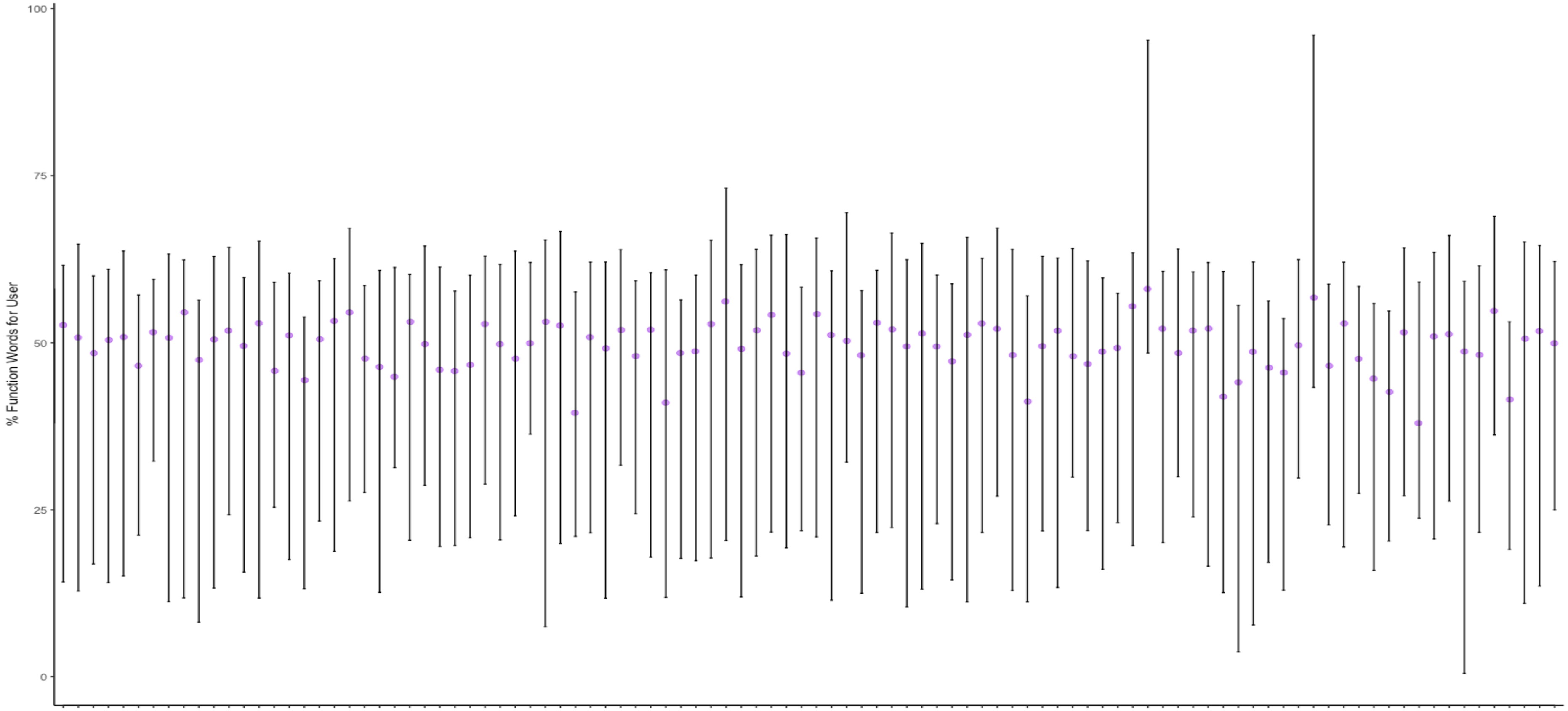
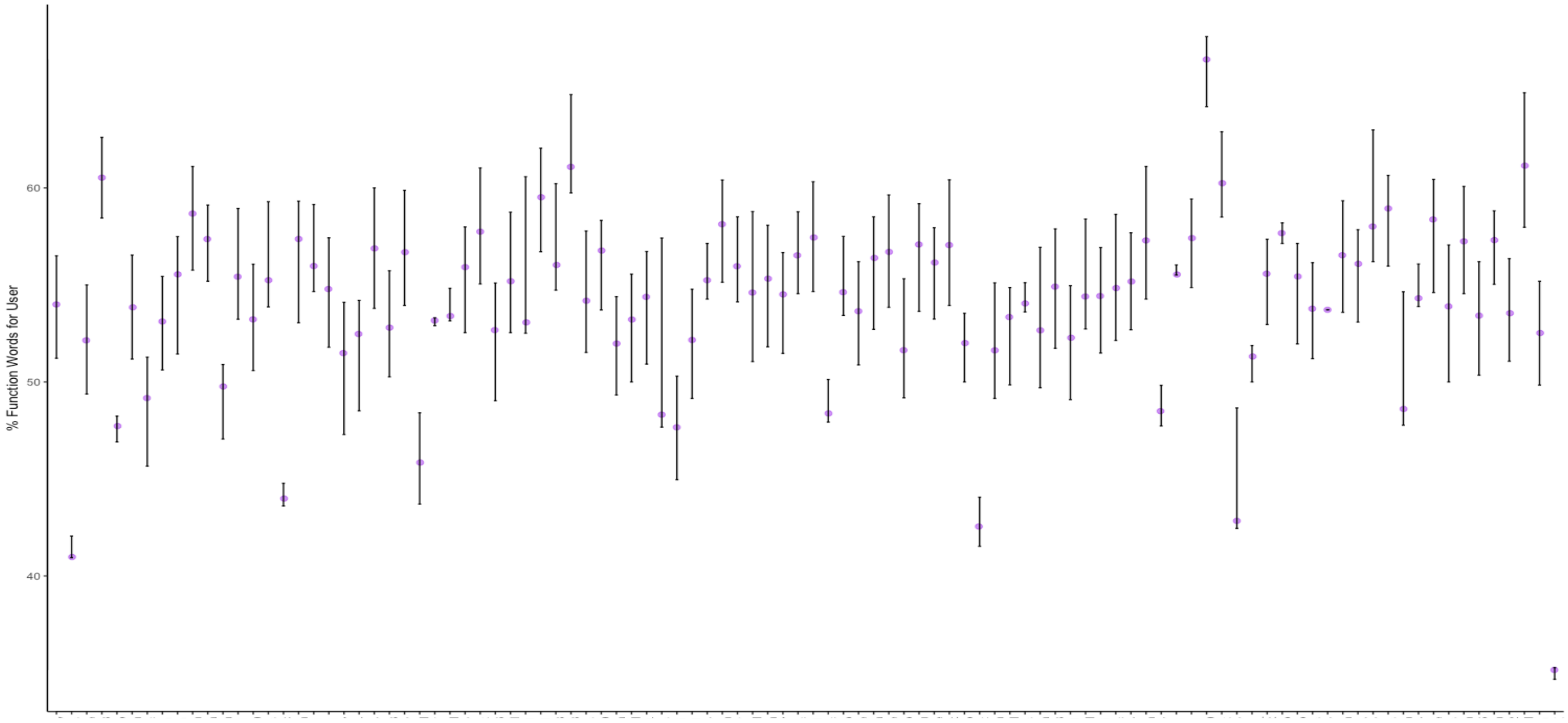


*Figure B. Top 100 users with the **lowest** variance of overall function word use (%). Each purple dot and line denotes a single unique user Variance was calculated based on all user's posts within the dataset (post-cleaning). The purple dot denotes the user's average number of function words across all of their posts. The error bar denotes the highest and lowest % function words across all posts. x-axis shows users.*



Figures A and B show that there is indeed much variance seen within the dataset in terms of the overall % of function word use. Similar patterns are seen for each specific type of function words (e.g., pronouns, articles, negations). Interestingly, in Figures A we note that some users' use of function words overall varied from ~5% to often ~60%, which is extremely wide.

In contrast, Figure B generally looks like most users here have ~10% variance in their use of function words. This early stage of findings indicates some level of individual differences, which may indeed be influenced by the social norms of the various communities' users are posting in. Therefore, these differences in communication behavior were deemed enough to warrant further data analysis interested in whether users do change and accommodate their linguistic style across communities (subreddits).

Individual Subreddit Linguistic Style (In)consistency

Similarly, we calculated the z-scores for each subreddit (N= 2,724) in order to examine the number of subreddits that are regarded as statistically extreme (>1.96 standard deviations from the mean) (Table B).

Table B. Z-scores of function words for each subreddit (N = 2,724). We focus only on the frequency of subreddits that fall >1.96 sd from the mean (95% confidence interval) and 3sd from the mean to show more 'extremely' inconsistent subreddits. We only look at inconsistence (e.g., >1.96 or 3 sd from the mean, rather than extremely consistent users <1.96 or 3).

Word Type	>1.96 sd from *M*	% of Users	>3 sd from *M*	% of Users
Function Words (Overall)	110	.040	47	.017
Personal Pronouns	111	.041	27	.010
Impersonal Pronouns	98	.036	33	.012
Articles	111	.041	35	.013
Prepositions	104	.038	39	.014
Auxiliary Verbs	100	.037	36	.013
Conjunctions	102	.037	25	.009
Negations	84	.031	30	.011

High Frequency Adverbs	106	.039	41	.015
Quantifiers	104	.038	41	.015
Number of Individual Users	2,724			

We find that again, subreddits are remarkably consistent and few subreddits appear to fall >1.96 standard deviations from the average variance found for all the subreddits included. However, here we must note that this is from the same dataset sliced in a different way (e.g., group versus individual) and therefore, it is largely to be expected that these results are extremely similar to results reported in Table B.

Appendix 3. Additional Analyses of Individual Types of Function Words Across Subreddits

The results reported in *Chapter IV* consider all nine dimensions of function words as an average. I demonstrated that users significantly shifted their linguistic style to match subreddits. However, I was additionally interested to test whether this holds when you consider each of the types of function words *separately*. The main interest here is to examine whether any function word use was *insignificant* as this might suggest that certain types of word do not shift with context changes. This could offer insight into elements of behavior or identifiers that remain stable for an individual, which may be useful within security research.

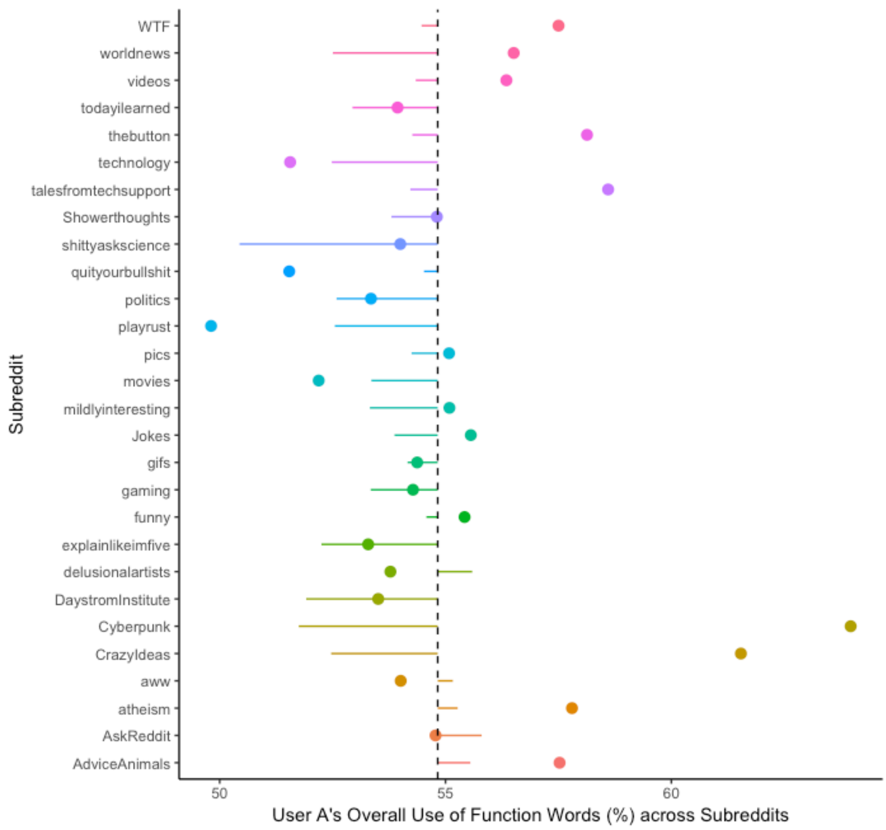
I therefore ran additional within-subjects ANOVAs in order to test whether individuals significantly shifted in terms of each dimension of function words in the five selected subreddits (r/AskReddit, r/politics, r/worldnews, r/funny, and r/atheism). Upon running within-subjects ANOVAs, I did find that three of the nine dimensions of function words were indeed insignificant: prepositions, $F(4, 3912) = .57$, $p = .68$; auxiliary verbs, $F(4, 3912) = 1.67$, $p = .15$; and finally, quantifiers, $F(4, 3912) = 1.87$, $p = .11$.

This is interesting both theoretically and in applied settings. This provides some evidence that types of function words are stable, aligning with prior literature (e.g., Pennebaker, 2011), which could be theoretically viewed as an element of one's identity that remains unchanged with context. Therefore, those interested in what elements of behavior and perhaps identity do not change and remain stable across contexts based on linguistic style, may wish to consider having some additional focus on prepositions, auxiliary verbs, and quantifiers. This additionally may help with data linking and matching users across various accounts. However, whether this could be used as any form of predictor or personal identifier required additional analysis. This is of course an early stage of research; however, it perhaps provides an additional starting point of interest.

Appendix 4. An Alternative Visualization of Individual User Variance and Communication Adaptation of Linguistic Style across Subreddits

As shown in *Chapter IV*, people vary in terms of their linguistic style. The additional analysis in Appendix 3 highlights that some function words are more stable than others across subreddits. This appendix provides an alternative visualization of an individual's use of function words. This aims to demonstrates the dynamic nature of linguistic style within some individuals (Figure C).

Figure C. An alternative way to visualize single individual's linguistic shifting across subreddits. This shows overall function word (%) usage. Colored lines denote average % of function words on the subreddit (the group). Each colored dot is the % of function words based on the users posts in that specific subreddit (the individual). The black dashed line denotes the individual's average use of function words across all posts.



While this thesis will be placed online, the offline version has an additional nine pages printed on acetate in order to provide a dynamic overview of one individual from the dataset used in *Chapter IV*. They were chosen as they had posted the most within this dataset (N > 2,000) and across many subreddits (N > 25). For privacy reasons their username is not revealed across the set of graphics.

The way data is visualized is critically important for conveying complicated information to readers. As data sets become larger and more complex, hence, the use of data visualization techniques is important for picking up errors and assessing data quality (e.g., missing data, data entry problems) (Heer & Kandel, 2012). Additionally, data visualization can be a powerful tool to convey meaning to expert and non-expert audiences as well as for teaching (Ellis & Merdian, 2015; Valero-Mora & Ledesma, 2014). As this chapter (and others) have used arguably large and complex data-sets, there has been a need to be creative in order to capture the constructs as required (e.g., Appendix 2). Dynamic data visualization via online applications such as Shiny in R is slowly becoming more common within psychological science (e.g., Ellis & Merdian, 2015). I intend to build a Shiny (or similar) app in line with the following pages, so online readers of the final publication will be able to see explore this dataset dynamically. However, as this thesis will firstly be printed, I wanted to capture some level of dynamic visualization *offline*. I am an artist that found academia, and throughout this thesis, I have greatly enjoyed the crossover between art and science, with data visualization arguably sitting between the two (Steele & Iliinsky, 2010).

The following nine graphs show the same user as reported in Figure C, but instead of looking at the overall use of function words, each graph looks at different dimensions (e.g., personal pronouns, impersonal pronouns, articles, etc.). Each graph can be viewed separately (e.g., only looking at prepositions, for example) by keeping the white A4 sheet of paper between each acetate page. However, you can also look at the individual's use of all function word dimensions across all 28 subreddits by removing all of the A4 white pages between the acetate sheets. Visualizations are stacked, which shows how dynamic and complex one person's linguistic style can be in multiple online communities hosted across Reddit.

CHAPTER V

SHAPE SHIFTING ACROSS SOCIAL MEDIA

<table>
<tr><th colspan="10">This declaration concerns the article entitled:</th></tr>
<tr><td colspan="10">Shape Shifting Across Social Media</td></tr>
<tr><th colspan="10">Publication status (tick one)</th></tr>
<tr><td>Draft manuscript</td><td></td><td>Submitted</td><td>X</td><td>In review</td><td></td><td>Accepted</td><td></td><td>Published</td><td></td></tr>
<tr><td>Publication details (reference)</td><td colspan="9">Davidson, B. I. & Joinson, A. N., [under review]. Shape Shifting Across Social Media.</td></tr>
<tr><th colspan="10">Copyright status (tick appropriate statement)</th></tr>
<tr><td colspan="3">I hold the copyright for this material</td><td>X</td><td colspan="5">Copyright is retained by the publisher, but I have been given permission to replicate the material here</td><td></td></tr>
<tr><td>Candidate’s contribution to the paper (detailed, and also given as a percentage).</td><td colspan="9">Formulation of ideas:
BID [70]; AJ [30]

Design of methodology:
BID [70]; AJ [30]

Analysis:
BID [100]

Presentation of data in journal format:
BID [80] wrote first drafts. AJ [20] edited subsequent iterations.</td></tr>
<tr><td>Statement from Candidate</td><td colspan="9">This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.</td></tr>
<tr><td>Signed</td><td colspan="5">Brittany Davidson</td><td>Date</td><td colspan="3">27th June 2019</td></tr>
</table>

In the previous chapters, I utilized a variety of methods and data sources in order to understand both what these types of data and methods can reveal about individuals or groups, but also in *Chapters III* and *IV* in particular, do users change and adapt themselves according to the context or over time? Reiterating that the key focus is understanding behavior – each chapter in this thesis utilized a form of digital trace (e.g., smartphone usage data, meta-data, content data) to capture actual usage and interaction with a technology rather than analyzing user experiences with technologies. This chapter addresses the first research question of this thesis in particular: how and why do people 'shape shift' across various online systems? It also provides an alternative approach to understanding user interactions with technologies, which provides answers to the second research question regarding what we can understand about people via different methods.

However, in contrast to all the previous empirical chapters (*II-IV*), this chapter takes an alternative approach to understanding user behavior. In this chapter, I take a qualitative approach using semi-structured interviews and the repertory grid technique. Here, this chapter seeks to demonstrate the importance of understanding user *experiences* with technology. These insights, while non-generalizable, begin to reveal *why* users may present and behave differently across different contexts. The findings show that actually there are a variety of reasons for 'shape shifting' behavior adaptation, which is important for applied research (e.g., marketing/advertising messages, security). This is discussed throughout this chapter.

Leading on from *Chapter IV,* we found that some users can, and do, adapt their linguistic style across online communities. We additionally showed that social media is a diverse experience for many users. There are a number of reasons and explanations as to *why* users 'shape shift', which is more specifically considered here. This is perhaps because individuals have distinct audiences across assorted services, and, perhaps, manage multiple personal and organizational identities on those services. However, while individuals may be able to manage audiences and maintain a separate work-life balance, this is becoming increasingly difficult online, especially in an age where social media platforms have become ever more interlinked and ubiquitous. Even once separate services (e.g., WhatsApp, Instagram) are now parts of a single organization (Facebook), often with shared authentication and access routes, merged content, and contacts suggested from one platform to the next. While this merging of multiple sites (with, potentially, different audiences) might not seem problematic in a world of a ‘single, authentic identity’ envisioned by Facebook’s CEO Zuckerberg, there is considerable evidence that a ‘single identity’ is neither natural nor usual in offline life. This exploratory study utilizes a novel and mixed methodological approach to better understand user behavior across social media platforms.

This study conducted 22 semi-structured interviews and employed a Repertory Grid Technique to reveal deeper similarities and differences in behavior across various online platforms. Drawing from social role and identity theory, we found that users had a variety of strategies for managing *multiple audiences* across *multiple platforms*. This most commonly occurred on Twitter and Instagram. Almost all participants actively separated their professional (e.g., LinkedIn) and socially (e.g., Facebook or Instagram) focused platforms.

Abstract

Individuals change and adapt their behavior according to their social situation (e.g., transitioning from work to home). However, how does this shape shifting of self-presentations and identity translate into various online platforms? This exploratory study utilizes a novel and mixed methodological approach to better understand user behavior across social media platforms. We interviewed 22 participants and employed a Repertory Grid Technique to reveal deeper similarities and differences in behavior across various online platforms. Drawing from social role and identity theory, we found that users had a variety of strategies for managing *multiple audiences* across *multiple platforms*. This most commonly occurred on Twitter and Instagram. Almost all participants actively separated their professional (e.g., LinkedIn) and socially (e.g., Facebook or Instagram) focused platforms. Implications for the planned merger of Facebook Messenger, Instagram, and WhatsApp are discussed.

1 Introduction

'You have one identity. […] The days of you having a different image for your work friends or co-workers and for the other people you know are probably coming to an end pretty quickly.' (*Mark Zuckerberg, as reported by (Zimmer, 2010))*

For many users, social media is a diverse experience. They have distinct audiences across assorted services, and, perhaps, manage multiple personal and organizational identities on those same sites and applications. However, while most individuals usually able to easily manage diverse audiences and maintain a separate work-life balance, this is becoming increasingly difficult online, especially in an age where social media platforms have become ever more interlinked and ubiquitous. Even once separate services (e.g., WhatsApp, Instagram) are now parts of a single organization (Facebook), often with shared authentication and access routes, merged content, and contacts suggested from one platform to the next. Indeed, in early 2019 Facebook announced plans to merge communication across Facebook Messenger, Instagram and WhatsApp (BBC, 2019). While this merging of multiple sites (with, potentially, different audiences) might not seem problematic in a world of a 'single, authentic identity' envisioned by Facebook's CEO Zuckerberg, there is considerable evidence that a 'single identity' is neither natural nor usual in offline life.

For instance, there has been extensive prior research into the 'multiple audience problem' posed by social media (e.g., Colliander et al., 2017; Marwick & boyd, 2011). Studies of single social media sites (usually Facebook) confirm that users have friends, colleagues, potentially bosses, and family connected to their profile as 'friends', which can cause anxiety and discomfort due to the discrepancies in term audience expectations about who we are, and how we should behave (e.g., Marder, Joinson, & Shankar, 2012; Rui & Stefanone, 2013; van Dijck, 2013). The presence of multiple audiences on social media sites means that users have to actively monitor their self-presentation in order to meet the different expectations of diverse groups, potentially leading to social anxiety, problems with social relations (Binder, Howes, & Sutcliffe, 2009). Presenting multiple facets of ourselves is not well supported on most single services, in part because group systems are under-utilized. As a consequence, many users report presenting to the 'lowest common denominator' where the most easily offended audience acts to 'chill' expression (Marder, Joinson, Shankar, & Houghton, 2016), meaning that online audiences to act as a type of information control (Hogan, 2010). This 'context collapse'(boyd, 2007) can lead individuals to only share content and information that is deemed acceptable to the broadest audience within their network. As such, an individual may have a seemingly neutral Facebook profile with little information regarding their sexuality and sexual preferences as they are 'friends' with their boss, family and socially distant colleagues.

A less well researched possibility is that users adopt different social media sites for distinct audiences and purposes. Thus, perhaps Facebook becomes the location for family and friends to keep in touch, LinkedIn becomes the place to build professional networks, Twitter becomes the site of choice for topic-based arguments with strangers, and Instagram is used for subtle flirtation and cyber-stalking. Indeed, anecdotal evidence suggests that social media is gradually becoming atomized in such a way, with Snapchat becoming (perhaps temporarily) the site of choice for building relations with people known offline (Piwek & Joinson, 2016), LinkedIn advertising itself as the 'world's largest professional network', and a range of sites used for sexual expression and flirting (Albury, 2017). The plan announced by Facebook to unify messaging across its multiple platforms (BBC, 2019) may challenge this atomization by effectively forcing users to integrate multiple versions of themselves into a single, Facebook-based, identity.

While people's behavior on social media sites has been the topic of considerable research over the last 10 years (e.g., Asur & Huberman, 2010; Kietzmann, Hermkens, McCarthy, & Silvestre, 2011), rarely are transitions and differences between sites and services considered in terms of both the uses to which they are put, and the implications of that for people's online identities. For instance, Trepte (2015) discusses how social media services might have 'warm' and 'cold' affordances that enable (or discourage) intimate behavior. However, implicit within Trepte's distinction is that the goals and motives of users remain consistent, but are buffeted by the affordances of the site they are using. The goal of the present research is to begin exploring if this is the case, and what the implications of platform-convergence might be for users. More specifically, we explore how people's self-presentation behavior, and imagined audiences, change as they move from one social media service (e.g., Facebook) to another (e.g., LinkedIn, Instagram).

Humans are fundamentally social beings – something that has continued, and become accentuated – as the world has become increasingly digitized (e.g., Dellarocas, 2003; Loebbecke & Picot, 2015). A key part to this sociality is our ability to change and adapt our behavior in response to both internal (e.g., emotion, expectations) and external (e.g., audience, environment) factors (e.g., Herrmann, Jahnke, & Loser, 2004; Hogg, Terry, & White, 1995). This social flexibility extends to our behavior across contexts – we wouldn't expect a person to behave identically as s/he transits from work to home, or from home to a night out with friends. This type of context-based behavioral change is normal and expected by those around us (e.g., Fiske, 2010; Herrmann et al., 2004).

This may also be seen in an online context, where a user will behave differently as they move from one online system to another. For example, the same colleague may be inclined to present themselves differently on LinkedIn compared to on Facebook, while their Tinder profile may be more different again. Indeed, there is some limited evidence for this – for instance, Vasalou and Joinson (2009) found that users were likely to create a more attractive avatar for a dating profile and a more 'intellectual looking' avatar for an online gaming profile. This does not mean that online identity is in someway not 'authentic' or 'genuine', but rather that identity is itself dynamic and changing, with different elements drawn into use according to the context and ongoing interaction.

One approach to identity that supports this view is to consider it in terms of **social roles** (Fiske, 2010). Throughout our lifespans, we adopt, lose, and shift social roles as we move from daughter to pupil/student, colleague, wife, mother and so on. Social roles are structured patterns of behavior (Ang & Zaphiris, 2010) that have been used to understand meaningful interactions between individuals within a network or system (Welser et al., 2011). At the same time, we respond to others according to the role they are currently playing (Hogg et al., 1995), making social roles an inherently fluid, dynamic approach to understanding how our identity changes according to the context we find ourselves in at any one moment.

An alternative approach to understanding online **identity** reflects the quote by Mark Zuckerberg at the beginning of the article. While there are many approaches to identity that understand it as multifaceted and dynamic, there is a strand of identity theory that stresses the importance of authenticity and a core 'self' that survives across contexts. For instance, Erikson (1959) suggested that identity stems from a young person's development (e.g., their experiences, culture, and history), which creates internal self-consistency and develops a persistent external character. From this perspective, an overly fragmented self, lacking in internal self-consistency and suffering from *Zerrissenheit* (inner conflict) is problematic for individuals (Sheldon, Ryan, Rawsthorne, & Ilardi, 1997). Sheldon et al. (1997) argue that multiple identities and shifts in persona come at a cost to one's wellbeing for variety of reasons, for example, social role conflict, when an individual has one or more incompatible social roles, and therefore will adopt some form of coping mechanism to resolve this (Biddle, 1986). However, if we adopt the perspective that identity is something an individual *does* rather than who the individual *is* (van Zoonen, 2013), then being able to enact a range of identity performances could be seen as beneficial for users. From this perspective, identity becomes a resource to be deployed as part of a self-presentation process, allowing us to impress others or to integrate with various social groups (Gonzales & Hancock, 2008). Self-presentation strategies have been examined in various online contexts, for example, forums, online gaming contexts, dating websites, and social media platforms (e.g., Ellison, Heino, & Gibbs, 2006; Papacharissi, 2002; Rui & Stefanone, 2013; Vasalou & Joinson, 2009). The affordances of social media services potentially allows for users to not only be more creative and flamboyant in terms of their self-presentation online (Papacharissi, 2002), but also to use multiple sites as a methods for dividing audiences and avoiding context collapse.

However, we know little about how people negotiate the potential discrepancies in their identity performance across multiple sites and services, what the impact of shared (or distinct audiences) might be, and what the challenges of unifying systems might be. Certainly, it is conceivable that users embrace uniform self-presentation across multiple systems. However, given the differing system designs, goals, and audiences of sites, we suspect that this is unlikely to happen (Levina & Arriaga, 2014). Instead, we anticipate that users will have a variety of strategies for managing *multiple audiences* across *multiple sites* that reflects both the likely audience, the publicness of the behavior, use of real identity, and the purpose, and social norms and warm / cold affordances of the site.

In the present research, we explore how people negotiate self-presentation and social roles across multiple sites. In particular, we are interested in the balance of audience, norms and purpose of the service, designed aspects of the service in presentation, and the ways presentation might relate to social roles adopted by the individual, and potential conflicts with the idea of a single, 'authentic' identity.

2 Methods

Given the exploratory nature of the research, we began with semi-structured interviews regarding participant's use of social media and communication platforms. We then used a method from counselling psychology – the Repertory Grid Technique (Kelly, 1955) – in order to reveal key similarities and differences across multiple platforms. This technique has been used across a variety of contexts from marketing (e.g., Lemke, Clark, & Wilson, 2011), information systems (e.g., Tan & Hunter, 2002) to software engineering (e.g., Edwards, McDonald, & Michelle Young, 2009). For the analysis, we used Braun and Clarke's (2006) Thematic Analysis Technique and NVIVO to identify and cluster emerging themes. This method was used for several reasons: 1) it is widely used for qualitative research and is a well-known method across social sciences (Braun & Clarke, 2006; Nowell, Norris, White, & Nancy Moules, 2017); 2) It offers considerable freedom and flexibility, which allows for rich insight (Nowell et al., 2017); and 3) It is well suited to exploratory research where there is little existing research. This study was ethically approved by the University of Bath and was conducted in accordance with the ethical guidelines of the British Psychological Association (BPS). IRB approval was not necessary for this study.

2.1 Participants

Twenty-two participants (14 female, mean age = 28.22, range: 22 –39) participated. Most participants were current students at the University of Bath. They were recruited via word-of-mouth and snowballing. Out of the 22 participants, one participant did not provide output for the Repertory Grid Technique. The participants were mostly

European, with two Chinese students, and an American. All participants took part voluntarily.

2.2 Semi-structured Interview

Upon arrival, we provided each participant with a series of cards with the names of various social networking sites (SNSs) and social media services (e.g., Facebook, Instagram, LinkedIn, Snapchat, WhatsApp, Twitter, Reddit, etc.); other cards were added if required. The participant was asked to give an overview of their online social media usage. They were then asked about why they still use platforms, why they subsequently deleted accounts, and about content shared on each account. Questions were asked about their audience and contacts, followers, or 'friends', and any concerns they had online. The interviews lasted approximately 40-50 minutes. The interviews were fully recorded and transcribed.

2.3 Repertory Grid Technique

Once the interview had finished, we moved onto the Repertory Grid Technique. There are three components of Repertory Grids (Kelly, 1955; Tan & Hunter, 2002):

1. *Elements* – the object or context of the research (typically people or versions of the self (e.g., ideal, real) in counselling settings, social media services here such as: Facebook; Instagram; and, LinkedIn)
2. *Constructs* – the participant's interpretation of how a selected triad of elements are similar or different (typically bi-polar such as 'private-public').
3. *Links* – the way in which all elements relate to constructs

Participants were asked to choose the top 8-10 social media platforms they used from a selection of flash cards. The cards were then shown in sets of three to the participant, who was then asked: '*Which two are similar, and by the same token, the third is different, and why?*'. The reason given for the similarity/difference was recorded as a 'construct'. As noted, constructs are bipolar in nature (Fransella, Bell, & Bannister, 2004; Kelly, 1955; Tan & Hunter, 2002), e.g., public-private, truthful-dishonest etc. Each construct is unique to each participant. We continued to cycle through these cards in unique triads until no new constructs were generated and began to repeat. Next, participants were asked to select their top five-used social media platforms. Then, we used the constructs previously elicited in order to score each of the five platforms in terms of each construct. For instance, if a construct was 'professionalunprofessional', the participant ranked each platform from the most professional to the least professional in terms of their usage. These scores were then used to calculate Euclidean Distances between the services for each participant (see next section for dendrograms of some participants). This process was also fully recorded and transcribed alongside the interviews.

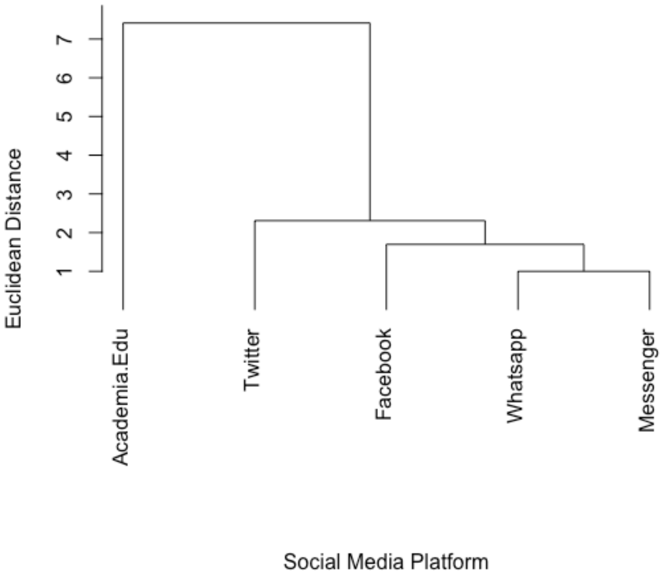
3 Finding and Discussion

In general, participants reported having different uses, purposes, and self-presentations across their various social media outlets. Overwhelmingly, most participants initially had a fairly binary approach, where they focused on the distinction between their professional and their social self-presentations. Unsurprisingly, platforms like LinkedIn and Twitter tended to be associated with work, while Facebook or Snapchat were viewed as social platforms. However, underneath the broad categorization of professional vs social platforms, there were subtle differences around audience management, conflicted self-presentation, and the ways in which the systems shape behavior. These themes are discussed in the next sections, referring back to the Repertory Grid outputs and interviews simultaneously.

3.1 Multiple Audiences, The Chill, and the Lowest Common Denominator

All but two participants reported that they attempted to maintain a separation between their professional and social oriented platforms. Most participants reported adapting their self-presentation across various social media platforms, with the exception of P10 and P18 who attempted to maintain a 'personal brand'. There was substantial agreement between participants on which platforms were associated with professional or social usage (e.g., *professional*: LinkedIn, Academia.edu; *social*: Facebook and Snapchat), which seemed to be reflective of a desire to manage different audiences, and the roles participants adopted in front of those same audiences. Unsurprisingly, Facebook was the most common social media platform for participants to have or had; it was also named as '*the most traditional'* social networking platform. Participants reported that their use of Facebook was primarily social and was often kept separate from professional platforms (see Figure 1).

Figure 1. P20's Euclidean Distance based on personal constructs from the interview. Graph shows the first split between Academia.Edu and Twitter, showing the professional vs social boundary. With the following social media platforms, Facebook, WhatsApp, and Messenger being similar in terms of her constructs, and were all regarded as social.



We can understand this almost binary approach to social media platforms – social or professional, by using the theory of the lowest common denominator (Hogan, 2010; Marwick & boyd, 2011). This considers how social media platforms are subject to multiple audiences, which can be difficult to satisfy. Therefore, this causes individuals to share content that will be the least offensive to the widest audience within the network (Hogan, 2010; Marwick & boyd, 2011) – also known as the 'chilling effect' (Marder et al., 2016). Hence, friends, followers, or contacts can be viewed as a form of information control and management (Hogan, 2010). This therefore creates an interesting dynamic between the individual and their audience due to the individual's awareness of potential surveillance from both peers and those in positions of power/authority (e.g. employers).

The professional-social separation – or even the management of various social spheres – is increasingly difficult online as many platforms continue to integrate (for example, Facebook's ownership of Instagram and WhatsApp). Within a single service, this has been called 'collapsed contexts', where individuals must handle multiple audiences simultaneously within a single platform (boyd & Ellison, 2007; Hogan, 2010; Marder et al., 2016; Marwick & boyd, 2011). For example, P10 reported that she misjudged her Instagram following/audience. Initially, her Instagram was used to share her makeup portfolio and later, she had started sharing posts of her in 'drag' makeup. This content offended several users, which led to her removing these posts and later deleting her account to avoid further negative feedback. This ability to remove accounts, as well as to create new ones in minutes, creates an interesting playing field in terms of playing new personae should one not suit the user (Lehdonvirta, 2010; Papacharissi, 2002).

Creating more conflict between and within platforms, social media sites are increasingly multipurpose and encourage users to share ever more elements of their life online (e.g., relationship statuses, work place, hobbies, etc.), which is not necessarily well-received (e.g., DiSalvo, 2010). For example, P19 cross-posted everything between Facebook and Instagram. She described her accounts as having, '*all of my photos and selfies*' as it is her way '*to express [her] feelings*'. This intense sharing behavior has been linked with increased wellbeing, where this sharing of information is therapeutic and aids emotion regulation especially after negative experiences (Buechel & Berger, 2012). However, other scholars suggest these 'updates' reflect narcissism and vanity (DiSalvo, 2010), which was reflected in some participant views on 'friends' posting heavily online. For example, P7 described her '*annoying friend*' that continues to post '*tens of baby photos every day*' - although, this did not deter her (and other participants) from remaining engaged online.

3.1.1 Virtual wall maintenance

It was clear that participants had distinctive self-presentations for different accounts. Most commonly, participants expressed a desire to maintain a separation between their professional and social life, which demonstrates a need for a certain level of social role consistency (Biddle, 1986) online. Unsurprisingly, participants often reported feeling uncomfortable when these boundaries were blurred. This aligned with Goffman's description of audience segregation breakdown leading to feelings of anxiousness (Goffman, 1956), reflecting recent literature examining multiple audience management and social anxiety (e.g., Marder et al., 2012):

> *'I always feel strange sending a message on WhatsApp to a supervisor or boss, as I might message and they're cooking dinner [...] it feels very personal, it doesn't feel quite right.' – P4.*

> '*I try not to have colleagues [on] Facebook because it is supposed to be a place to vent,' – P21.*

Several participants reported having highly restrictive privacy settings on Facebook, which was often discussed in light of efforts to maintain a separation between their professional and social lives. Privacy settings were typically used, however, to keep work-based contacts outside of the Facebook platform rather than to filter content to different groups within Facebook, which aligns with prior research reporting a lack of use of audience management tools (Marder et al., 2012). One participant, P6, was the only interviewee who utilized Facebook's privacy settings to manage audiences within the site – in her case limiting access to albums with photos of her daughter to family members. Other participants managed social-professional audience tensions by carefully managing their posting behavior – for instance, P20 reported that she '*used to be more provocative*' but now vets her posts since she has several colleagues as 'friends' and did not want to share anything '*embarrassing*' or '*too revealing*'. P1 was also wary of Facebook posts as they allowed for (often critical) feedback. He enjoyed Snapchat and saw it as distinctive because he enjoyed its flippant and disposable nature, whereas Facebook and WhatsApp were primarily for '*serious'* communication. This distinction reflects previous work on the value of playful (disposable) platforms – what we might think of as having 'warm' communication affordances – in building bonds between people (Piwek & Joinson, 2016).

3.2 Self-Presentation, Conflicts, & Identity Crises

3.2.1 Multiple Self-Presentations

Participants largely reported having social and professional accounts across social media. The majority of participants reported posting infrequently to Facebook, where they only did if they felt it was important and of interest to their 'friends'. For instance, P1 shared politically or environmentally focused posts to raise awareness, which he anticipated made his self-presentation on Facebook serious. P21 reported she only posted on Facebook if a major life event had happened – for example, her graduation. Similarly, while most participants had a LinkedIn account – they did not engage often or keep their account up-to-date: '*I am pretty half-assed maintain[ing] a profile, I don't really engage on there,'* [P4]; '*I have a profile, I have not updated it […], I feel like a bottle of shampoo marketing myself on there […], I think it's a lot of lies and untruths,*' [P6]. Beyond sharing work-histories, most participants did not engage or share additional information on their LinkedIn accounts. This infrequent posting discussed by several participants across many social media sites appeared to be a method to handle multiple audiences, where there was a consistent concern with how their content would be perceived by 'friends' or followers.

Self-regulation of content was also often discussed in relation to their social accounts in order to avoid conflict online or some self-regulating to maintain their physical image. For example, P12 expressed concern about photos being uploaded from nightclub Facebook accounts after a night-out, or 'friends' tagging photos of her looking '*ugly*'. Whereas P2 welcomed less edited and more realistic portrayals on Facebook, '*so you can go back and have memories [that] I remember,*' rather than '*looking back at a lot of photos that you've tried to make different from […] reality*'. This reflects the importance of self-presentation and the way users believe they are being perceived by their audience(s). It was clear there were subtle differences in self-presentation between Facebook and Instagram. For instance, while P12 used Instagram frequently and there was some overlap between Facebook and Instagram content, all photos from a night out or a holiday would be posted to Facebook, while Instagram only received the '*best photo*'. Typically, Instagram (more so than Facebook) was consistently stereotyped as a '*heavily edited version of your life*,' and a '*perfect world*' portrayal online. P4 maintained a different form of separation between her Instagram and Facebook posts, with Instagram focusing on food, drinks, and art, which she believed is more suitable for an Instagram audience. Her posts were only occasionally cross-posted to Facebook if 'friends' were tagged in photos. This demonstrates different self-presentations – or social roles – being created online that are often kept separate from one another, which tended to be rationalized by wanting to please or fulfil their 'audience' expectations. Additionally, this demonstrates that the notion of a single, authentic identity based on Facebook suggested by Mark Zuckerberg, as reported by (Zimmer, 2010) will be a challenge and remains arguably unnatural.

This focus by some users on the audience over authenticity was not well-received by all participants. Several participants claimed Instagram was about '*building up a fake lifestyle*', which made it difficult to distinguish actual (photography) skill from '*too many filters*' – where '*everyone believe[s] they're a professional photographer*'. Yet, the participants reporting this concern still had accounts – even if they were not frequent users. This may relate to the notion of the 'fear of missing out' (FOMO), which is defined by '*the desire to stay continually connected with what others are doing*,' (Przybylski, Koutamanis, DeHaan, & Gladwell, 2013, p. 1841), as these accounts – despite some reported infrequent use – allowed users to maintain a silent presence and connection to 'friends' or followers (akin the 'social stalking' discussed in early studies of Facebook use) (Ellison, Steinfield, & Lampe, 2007; Joinson, 2008). However, other participants reported that they would unfollow updates from people who they deemed as oversharing, suggesting some form of calculus between the value of passive social information consumption and the need to manage the over-production of information by selected people.

3.2.2 Audience Mis-Match

Self-presentation behaviors are critical for forming outside impressions, where social media allows us to emphasize and explore new facets of ourselves or to be someone entirely different – should we want to (Papacharissi, 2002; Vohs, Baumeister, & Ciarocco, 2005). However, when there are clear discrepancies (known to the user or not) – there can be audience mis-matches, where the user will receive potentially negative feedback due to these discrepancies. For example, P10 used Instagram to show her makeup artistry portfolio and as this developed over time, she began to experiment with 'drag' makeup, which was not well-received by her audience. This dampened her Instagram experience to the point that she deleted her account. This demonstrates that while one can creatively build their self-presentation, the audience is also a key component of this identity development (Belk, 2013). P10 arguably experienced a conflict between her wanted 'self-presentation' or 'social role' on Instagram and her audience's expectations of her. This can also be understood by considering Higgins Self-Discrepancy Theory (1987), which sought to theorize negative affect associated with discrepancies within self (e.g., not fulfilling your own expectations) or in relation to other's expectations of the individual. Here, P10's discrepancy caused her to subsequently modify her behavior to meet the expectations of her audience due to uncomfortableness from upsetting her followers. However, this micromanagement of her account became '*unnatural*', hence she decided it was easier to remove herself from the platform. It is also interesting to note that P10 was one of the two people who reported using all social media platforms for the same purpose, in effect having a 'personal brand'. Yet, she had also experimented with her self-presentation on Instagram, suggesting that the brand was not quite as complete as she suggested. This highlights difficulties of self-report via interviewing as there can be discrepancies between what is reported and actual behavior (e.g., Ellis, Davidson, Shaw, & Geyer, 2019).

Other participants took advantage of anonymity to avoid audience mis-matches. For instance, P3 used an anonymous account on DeviantArt for his personal poetry. He kept this poetry account entirely hidden due to his uncomfortableness of peer-to-peer surveillance, and he noted that he didn't want this personal content shared with his Facebook 'friends' or audiences. With similar concerns, P20 ran two Twitter accounts, one for work and her other personal account, where she tweeted about politics and engaged in Twitter arguments. These two accounts were also actively kept separate via different usernames. This indicates that participants actively created and maintained boundaries both between and within their social media accounts.

While it is typical for our Facebook account to reflect our social selves and Academia.edu a professional self, some participants reported that other platforms were not as straightforward to present themselves on. For example, P20 reported a lack of clarity as to how she should appear and market herself on LinkedIn, which caused her to disengage with LinkedIn. Other participants (e.g. P20) willingly shared their multi-faceted career paths on LinkedIn despite a belief that it might look like she was '*just confused or lying*'. We can conceptualize this identity confusion online with the notion of identity crisis (Erikson (1959)). Erikson's theory of identity crisis related originally to children and young adults and was caused by their having too many social role choices. It may be that the multiplicity of social media platforms and audiences has extended this overwhelming amount of choice to adults too. As online, users will have an unlimited amount of identities to choose form – the only limit is the individual's imagination. This could lead to additional feelings of anxiety or stress from maintaining or the ability to continually generate multiple social roles or personae online that may conflict with other social roles the individual has (Biddle, 1986; Erikson, 1959).

3.3 Systems Shaping our Identity, Self-Presentation, and Behavior

The presentation the self portrayed online across platforms begins undoubtedly with the individual. However, the audience also impacts the on-going development of the digitized self (Belk, 2013). This can lead to affirmative experiences, for example, P12 reports positively of her Facebook 'friends' tagging, uploading photos, and commenting on each other's' profiles, or negative experiences, like P10 and the closing of her Instagram account. The audience's impact on identity can be understood via social role theory as the individual plays their social role in regard to the roles of the audience around them (Herrmann et al., 2004), or simply in regard to the individual wanting to abide by the social norms of the platform itself via social learning theory (Bandura, 1971). However, several participants stated how their posting habits varied across platform despite there being an overlap in 'friends', followers, or contacts between platforms. This therefore suggests that considering the audience alone does not capture entirely why user behavior will vary across even similar types of social media platforms. We therefore consider how platform itself plays a role in shaping the way in which the user will behave within this system (Levina & Arriaga, 2014).

Some participants mentioned they had overlap in Facebook 'friends' and Instagram followers and yet they share different posts and information on each site. Design-wise, there are several likely reasons for this. Firstly, Facebook privacy settings in our sample tended to be set to be reasonably private, whereas Instagram accounts were almost entirely public. Therefore, naturally there is a security question on what information you share with an often smaller and more closed network of people versus an entirely open and public profile, aligning with Gonzales and Hancock's (2008) distinction between seemingly public versus private platforms. Hence, the way users allow others' access to their account is important. On Facebook a granted 'friend' request is often required to gain access to someone's profile, whereas on Instagram users can simply follow anyone – unless individuals go out of their way to set their account to private, which was uncommon in our participants. This again sets a different tone with how users will interact with each other, for example, P18 has stated she was comfortable with the open nature of Twitter followers and not knowing who is following her tweets, however, when this translates to Facebook, she refuses unknown 'friend' requests. Interestingly again, P18 was the other participant that reported having a '*personal brand*' across all social media accounts and cross-posting content. This further demonstrates that those who are attempting to use social media for a brand, whether a person or company, that the social norms of the group, the system itself, and the content, is critically important to engagement and interaction.

Additionally, the way the system is designed creates a type of community – for example, Facebook encourages users to share all aspects of their lives via 'stories', live updates, status updates, photo and video sharing, reflecting Mark Zuckerberg's '*you have one identity*,', whereas Twitter or LinkedIn does not have all of those capabilities, which will impact what users do, purely based on the functionality of the system. Further, Facebook uses the term 'friends' for contacts, whereas Instagram and Twitter have 'followers' that can be asymmetric, potentially emphasizing that these sites are more for sharing and disseminating information or content and having access to vast amounts of information, often via hashtag searching. Additionally, other affordances differ between platforms, for example, Twitter has restricted character limits, unlike Facebook, which will likely to impact on communication behaviors (Wall, Taylor, Dixon, Conchie, & Ellis, 2013).

Some social media is more customizable than others. For instance, LinkedIn was largely regarded as inflexible and hard to personalize. In contrast, Facebook allows users to customize their profile image/video, cover image, biography, and newer updates such as 'featured photos'. Personal websites allow individuals to have complete control over their self-presentation – whereas, LinkedIn only allowed users to change profile images and the information they share. There is less capability to customize the overall look of one's profile to portray more personality via image or video. We anticipate this lack of engagement with LinkedIn perhaps relates to the lack of change within the system itself, where the ability to customize lacks in comparison to other systems. This aligns with Schmader and Sedikides (2018), where they suggest that the way an individual 'fits' with an online community or system is critically important to their overall engagement. We anticipate this lack of 'fitting' or ability to portray themselves or intended the role(s) effectively may constitute to low engagement.

4 Conclusions and Future Research Directions

This article intended to extend knowledge regarding individual's identity and self-presentation negotiation across various social media platforms. We anticipated participants would have various identities and self-presentations online. Hence, we expected participants would have several strategies in order to manage *multiple audiences* across *multiple platforms* that reflects both their audience – of 'friends', followers, or contacts, the level of privacy they maintain for their profile, use of their real identity (e.g., using their actual name), the purpose of the profile, and social norms naturally residing in the site.

Almost all participants separated their professional and social selves across various social media platforms. This shows that to a certain extent, all participants did maintain several self-presentations across multiple platforms. In terms of professional platforms, naturally, LinkedIn, Academia.edu, and often Twitter were commonly associated. Participants shared less intimate and detailed information as these platforms are almost exclusively public. This self-regulating behavior typically aligns with Bazarova and Choi (2014) and Gonzales and Hancock's (2008) distinctions between information shared across public and private online platforms. Participants engaged in some level of self-regulation behaviors for their social platforms, and it was found participants often shared different content across each site. Most commonly, this distinction was between Facebook and Instagram, where Facebook was often regarded more realistic and natural. In comparison, Instagram tended to receive the most polished images and content shared was more 'artsy'. This further aligns with the theory of the lowest common denominator (Hogan, 2010) and the notion of 'collapsed contexts' (boyd & Ellison, 2007). While there are tools inbuilt to restrict friend/contact access to posts, these tools were vastly underutilized, aligning with Marder et al. (2012). Participants preferred to de-tag or deleting old content, refrain from posting, or deliberately refuse 'friend' or contact requests from colleagues or bosses, most commonly on Facebook. This further shows that a single online identity is arguably not viable, and this may lead to user issues regarding to self-presentation across online contexts.

Having multiple narratives of self, or (subtly) different social roles, offline can be a source of role conflict and identity inconsistency, which can be a cause for concern and become stressful to the individual (e.g., Biddle, 1986; Erikson, 1959). From our findings, we argue that this almost flippant ability to 'try on' and experiment with new identities online could be a problem if the level of anonymity is lower (where 'friends', followers, or contacts know the user) and this creates a discrepancy in audience expectations. We report that audience(s) within social media sites are important for user self-presentation and identity development (Belk, 2013) and from our sample, if there are large discrepancies, this can cause user disengagement with the platform.

Finally, we consider how the systems themselves shape user-behavior and the audience that comes with these systems. For instance, while several participants had active Facebook and Instagram accounts, despite overlap in ‘friends’ and followers, posts and content shared on each site was tailored and adapted. This shows there are subtle differences between self-presentations and platform – which indicates audience alone is not sufficient to explain variance in online self-presentations. We found that the purpose, the ability to customize profiles, and the innovativeness of the platform all were important in determining participant usage and engagement. Additionally, how users connect with others is important. For instance, Instagram or Twitter are automatically public until these settings are actively changed. Whereas Facebook requires a ‘friend’ request in order to access a user’s profile information. From the outset, this provides a barrier that other platforms do not have, hence it was common for participants to maintain different privacy settings and contact restrictions with different sites.

This paper sought to provide a novel approach to explore online settings by utilizing Repertory Grids alongside semi-structured interviews, specifically here, identity and self-presentation across platforms. We found that participants do maintain multiple presentations of self across multiple sites, and do engage in a set of (not necessarily efficient) self-regulating behaviors in order to avoid offending others (Hogan, 2010; Marder et al., 2016). We find that there are many factors that impact participant behavior and engagement online. Understanding online behavior is complicated and clearly changes from platform to platform and while we found overlap between our participants; their individual usage and reasoning behind it, was unique.

References

Albury, K. (2017). Sexual Expression in Social Media. *The SAGE Handbook of Social Media*. https://doi.org/10.4135/9781473984066.n25

Ang, C. S., & Zaphiris, P. (2010). Social Roles of Players in Mmorpg Guilds. *Information, Communication & Society*, *13*(4), 592–614. https://doi.org/10.1080/13691180903266952

Asur, S., & Huberman, B. A. (2010). Predicting the Future with Social Media. In *In Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. https://doi.org/10.1109/WI-IAT.2010.63

Bandura, A. (1971). Social learning theory. *Social Learning Theory*. https://doi.org/10.1111/j.1460-2466.1978.tb01621.x

Bazarova, N. N., & Choi, Y. H. (2014). Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of Communication*, *64*(4), 635–657. https://doi.org/10.1111/jcom.12106

BBC. (2019). Facebook to integrate WhatsApp, Instagram and Messenger. Retrieved April 16, 2019, from https://www.bbc.co.uk/news/technology-47001460

Belk, R. W. (2013). Extended Self in a Digital World. *Journal of Consumer Research*, *40*(3), 477–500. https://doi.org/10.1086/671052

Biddle, B. J. (1986). Recent Development in Role Theory. *Annual Review of Sociology*, *12*, 67–92.

Binder, J., Howes, A., & Sutcliffe, A. (2009). The problem of conflicting social spheres: effects of network structure on experienced tension in social network sites. In *In Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 965–974).

boyd, danah, & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*, 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buechel, E., & Berger, J. (2012). Facebook Therapy? Why People Share Self-Relevant Content Online. *Advances in Consumer Research*, *40*, 203–208. Retrieved from http://www.acrwebsite.org/volumes/1012611/volumes/v40/NA-40

Colliander, J., Marder, B., Falkman, L., Madestam, J., Modig, E., & Sagfossen, S. (2017). The social media balancing act: Testing the use of a balanced self-presentation strategy for politicians using twitter. *Computers in Human Behavior*, *744*, 277–285.

Dellarocas, C. (2003). The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science*, *49*(10), 1407–1424. https://doi.org/10.1287/mnsc.49.10.1407.17308

DiSalvo, D. (2010). Are Social Networks Messing with your Head? *Scientific American Mind*, *20*(7), 48–55.

Edwards, H. M., McDonald, S., & Michelle Young, S. (2009). The repertory grid technique: Its place in empirical software engineering research. *Information and Software Technology*. https://doi.org/10.1016/j.infsof.2008.08.008

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? [In Press]. *International Journal of Human-Computer Studies: Special Issue on Human Accuracy*, 12. https://doi.org/10.31234/osf.io/6fjr7

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Ellison, N., Heino, R. D., & Gibbs, J. (2006). Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment. *Journal of Computer-Mediated Communication*, *11*, 415–441. https://doi.org/10.1111/j.1083-6101.2006.00020.x

Erikson, E. (1959). *Identity and the Life Cycle*. London, UK: W. W. Norton & Company Ltd.

Fiske, S. T. (2010). *Social Beings* (2nd ed.). NJ, USA: Wiley.

Fransella, F., Bell, R., & Bannister, D. (2004). *A Manual for Repertory Grid Technique* (2nd Editio). Chichester, England, UK: John Wiley & Sons Ltd.

Goffman, E. (1956). Embarrassment and Social Organization. *American Journal of Sociology*, *62*(3), 264–271.

Gonzales, A. L., & Hancock, J. T. (2008). Identity Shift in Computer-Mediated Environments. *Media Psychology*, *11*(2), 167–185. https://doi.org/10.1080/15213260802023433

Herrmann, T., Jahnke, I., & Loser, K. U. (2004). The Role Concept as a Basis for Designing Community Systems. *Cooperative Systems Design, Scenario-Based Design of Collaborative Systems*, 163–178.

Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological Review*, *94*(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, *30*(6), 377–386. https://doi.org/10.1177/0270467610385893

Hogg, M., Terry, D., & White, K. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*. https://doi.org/10.2307/2787127

Joinson, A. N. (2008). Looking at, looking up or keeping up with people?: motives and use of facebook. In *Proceedings of the SIGCHI conference on Human Factors in Computing* (pp. 1027–1036).

Kelly, G. (1955). *The Psychology of Personal Constructs*. New York, USA: W. W. Norton & Company Ltd.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241–251.

Lehdonvirta, V. (2010). Virtual Worlds Don't Exist: Questioning the Dichotomous Approach in MMO Studies. *The International Journal of Computer Game Research*, *10*(1). Retrieved from http://gamestudies.org/1001/articles/lehdonvirta

Lemke, F., Clark, M., & Wilson, H. (2011). Customer experience quality: An exploration in business and consumer contexts using repertory grid technique. *Journal of the Academy of Marketing Science*. https://doi.org/10.1007/s11747-010-0219-0

Levina, N., & Arriaga, M. (2014). Distinction and Status Production on User-Generated Content Platforms: Using Bourdieu's Theory of Cultural Production to Understand Social Dynamics in Online Fields. *Information Systems Research*, *25*(3), 468–488.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems*, *24*, 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Marder, B., Joinson, A., & Shankar, A. (2012). Every post you make, every pic you take, I'll be watching you: Behind social spheres on facebook. *Proceedings of the Annual Hawaii International Conference on System Sciences*, (July 2015), 859–868. https://doi.org/10.1109/HICSS.2012.12

Marder, B., Joinson, A., Shankar, A., & Houghton, D. J. (2016). The Extended "Chilling" Effect ofFacebook: The Cold Reality of Ubiquitous Social Networking. *Computers in Human Behavior*, *60*, 582–592. Retrieved from http://opus.bath.ac.uk/49558/1/Chilling_Effect_CHB_revised_submission_BM.pdf

Marwick, A. E., & boyd, danah. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media and Society*, *13*(1), 114–133. https://doi.org/10.1177/1461444810365313

Nowell, L. S., Norris, J. M., White, D. E., & Nancy J. Moules. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1). https://doi.org/10.1177/1609406917733847

Papacharissi, Z. (2002). Presentation of Self In Virtual Life: Characteristics of Personal Home Pages. *Journalism & Mass Communication Quarterly*, 643–660.

Piwek, L., & Joinson, A. (2016). "What do they snapchat about?" Patterns of use in time-limited instant messaging service. *Computers in Human Behavior*, *54*(January), 358–367. https://doi.org/10.1016/j.chb.2015.08.026

Przybylski, A. K., Koutamanis, M., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, *29*(4), 1841–1848.

Rui, J., & Stefanone, M. A. (2013). Strategic self-presentation online : A cross-cultural study. *Computers in Human Behavior*, *29*(1), 110–118. https://doi.org/10.1016/j.chb.2012.07.022

Schmader, T., & Sedikides, C. (2018). State Authenticity as Fit to Environment: The Implications of Social Identity for Fit, Authenticity, and Self-Segregation. *Personality and Social Psychology Review*, *22*(3), 228–259. https://doi.org/10.1177/1088868317734080

Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait self and true self: Cross-role variation in the Big-Five personality traits and its relations with psychological authenticity and subjective well-being. *Journal of*

Personality and Social Psychology, *73*(6), 1380–1393. https://doi.org/10.1037/0022-3514.73.6.1380

Tan, F. B., & Hunter, M. G. (2002). The Repertory Grid Technique: A Method for the Study of Cognition in Information. *MIS Quarterly*, *26*(1), 39–57. https://doi.org/10.2307/4132321

Trepte, S. (2015). Social Media, Privacy, and Self-Disclosure: The Turbulence Caused by Social Media's Affordances. *Social Media + Society*, *1*(1). https://doi.org/10.1177/2056305115578681

van Dijck, J. (2013). "You have one identity": performing the self on Facebook and LinkedIn. *Media, Culture & Society*, *35*(2), 199–215. https://doi.org/10.1177/0163443712468605

van Zoonen, L. (2013). From identity to identification: Fixating the fragmented self. *Media, Culture and Society*, *35*(1), 44–51. https://doi.org/10.1177/0163443712464557

Vasalou, A., & Joinson, A. N. (2009). Me, myself and I: The role of interactional context on self-presentation through avatars. *Computers in Human Behavior*, *25*(2), 510–520. https://doi.org/10.1016/j.chb.2008.11.007

Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-Regulation and Self-Presentation: Regulatory Resource Depletion Impairs Impression Management and Effortful Self-Presentation Depletes Regulatory Resources. *Journal of Personality and Social Psychology*, *88*(4), 632–657. https://doi.org/10.1037/0022-3514.88.4.632

Wall, H. J., Taylor, P. J., Dixon, J., Conchie, S. M., & Ellis, D. A. (2013). Rich contexts do not always enrich the accuracy of personality judgments. *Journal of Experimental Social Psychology*, *49*(6), 1190–1195.

Welser, H., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. *Proceedings of the 2011 IConference on - IConference '11*, 122–129. https://doi.org/10.1145/1940761.1940778

Zimmer, M. (2010). Facebook's Zuckerberg: "Having two identities for yourself is an example of a lack of integrity." Retrieved January 30, 2018, from http://www.michaelzimmer.org/2010/05/14/facebooks-zuckerberg-having-two-identities-for-yourself-is-an-example-of-a-lack-of-integrity/

CHAPTER VI

SUMMARY AND FUTURE RESEARCH

1 Thesis Overview

This 'chapter' acts as a final summary and overview of the thesis, noting that each chapter already touches on limitations, future work, and conclusions more specific to the paper itself. Here, I first provide some final thoughts reflecting back on the collective thesis before discussing potential ways forward and future research directions.

This thesis sought to investigate whether identity and behavior may change and adapt across online systems. It utilized a variety of theoretical frameworks, data sources, and methods in order to better understand individual and group-level interactions with technology between different contexts, devices, and over time. As discussed in *Chapter I* social psychology has increasingly relied on self-report instruments at the expense of direct behavioral measurement (Sassenberg & Ditrich, 2019). In response, I considered what different approaches, methods, and data sources can reveal about individual and group interactions with technology.

Each chapter (bar *Chapter V*) analyzed a different 'digital trace', from smartphone usage data (*Chapter II*), to meta-data (*Chapter III*), and textual content data (*Chapter IV*) in order to understand user behavior. This demonstrates primarily that online behavior is complex, dynamic, and noisy. Hence, we should not reduce this complex behavior to a 'digital' or binary interaction – where an interaction with technology is 'good' or 'bad', which is typically found in social psychological research (see *Chapter II*). We realistically have much to learn about identity and behavior online and how or why this may change in different settings prior to focusing on often negative outcomes.

While behavioral changes seen in *Chapters III* and *IV* can be partly explained by social psychological theories like *Social Role Theory* or *Identity Theory*, the ability to predict changes in behavior or what someone may do (or not do) next remains difficult. This is particularly pertinent for those interested in the link between online and offline behavior, all of which has tremendous implications for security practitioners (Katos & Bednar, 2008). Therefore, understanding objective interactions and behavior in different settings from objective data sources is critically important both theoretically for psychology as well as to a variety of applied settings.

Psychology must start utilizing opportunities from technologies (e.g., computational techniques for data analysis, scraping data, using smartphones or other devices), as there is an abundance of data sources and ways to understand individual and group behavior. While utilizing (often) large datasets that will naturally be noisy and require substantial cleaning, the objective and sometimes real-time data it produces is highly valuable to social psychology. Psychology needs to adapt and adopt new perspectives and approaches to remain relevant in an increasingly digitized society.

1.1 Key Research Questions Revisited

The goal of the research in the present thesis was to develop and apply new methodological approaches when studying peoples' behavior via technology. This enables us to gain new insights about how people behave when interacting using various technologies – from specific services (e.g., social media) to devices themselves (e.g., smartphones). Hence, the initial focus of the research presented in this thesis explored changes in behavior across various online or digital contexts:

> *RQ1: Do individuals adapt their behavior across different systems or over time? If so, how can this be measured and theorized?*

Yes – across this thesis, I have documented how individuals and groups adapt their behavior across different systems (*Chapters IV* and *V*) and over time (*Chapter III*). This was observed from both objective behavioral measures as well as reported experiences with technologies.

Many of these findings can be considered within framework afforded by *social role theory*. Here it is argued that individuals adopt, shift, gain, or lose various social roles across the lifespan (Fiske, 2010). For example, it is possible for a person to transfer from being an undergraduate student, to completing a postgraduate degree, and then a job. These roles can exist alongside being a partner, sister, and daughter. Our role composition and configuration is ever-changing, and this thesis has demonstrated how social roles shift when people are using a variety of online services (see *Chapters III, IV* and *V)*. For example, in *Chapter III*, we used both *social role theory* and *the reader-to-leader* framework to test whether social roles are indeed present in online communities, which we did indeed find. In addition, we mapped these social roles to the *reader-to-leader framework* (Preece & Schneiderman, 2009) to examine leadership within online ideological communities. When examining users over time, we found that they do change roles over time, which allowed us to map the most common role change pathways to demonstrate how diverse and non-linear behavior change can be, which aligns with Preece and Schneiderman's (2009) framework. This shows that behavior and elements of identity are indeed active and dynamic.

This active and dynamic nature of behavior was also demonstrated in *Chapter IV*, where we argued that if users adapt their linguistic style (LS) to match that of the specific community, this may reflect a social role change or shift if users consistently matched community LS. Hence, this chapter considered to what extent users converged (or diverged) their linguistic style to match (or not) various online communities. In line with previous research, we found that many users actually diverged from the community linguistic style (Jones, Cotterill, Dewdney, Muir, & Joinson, 2012). We did find that some users did converge towards several subreddits consistently, suggesting various levels of chameleon-like behavior by matching community linguistic norms. This provides a starting point where other types of linguistic analysis (e.g., NLP (pattern analysis), topic modelling) may be useful to extend and develop this work to demonstrate whether (or not) people will adapt their communication style across contexts as they shift roles. Additionally, we found that active moderation of communities significantly impacted user behavior with users diverging more when there was little to no content moderation. Hence, audience and social norms of communities changes behavior in line with Matias (2019). *Chapter V* took an alternative and arguably more traditional approach in psychological science. I interviewed and employed repertory grids to better understand experiences of behavior, self-presentation, and identity across social media. It was clear that users did actively manage and adapt their behavior and identities across multiple platforms. This aligns with social role theory and the prior quantitative chapters.

However, alternatively, we can also consider many of these findings in the context of identity theory, which argues that there is a 'core' self that survives across context. Van Zoonen (2013) suggests that identity is something we *do* rather than who we *are*, which allows for identity to be a core element of self that is employed to build various self-presentations, which is another perspective on the general findings of this thesis. Although, the Appendix 4 of *Chapter IV* demonstrates that there are certain function words (within our LS) that stay stable across contexts. This could be one way to quantify more stable elements of identity, which aligns with a more traditional perspective on identity and the 'core' self that remains stable (Erikson, 1959). This could suggest an exploratory new research avenue for quantifying 'identity'. However, this would require the development of new methods to potentially adapt or inform new theory for digitized social sciences. This leads into the second objective of this thesis:

> *RQ2: What can we understand about an individual's interactions with technologies from a variety of approaches, data sources (e.g., usage, meta-data), and methods?*

By utilizing various data sources, we can infer how they use technologies (e.g., length of time (*Chapter II*)), who they communicate with (via usernames online – *Chapter III*), patterns of behavior and communication (*Chapter III* and *IV*). This data can be used to understand groups or individuals, which will have a variety of applications from theorizing the underlying mechanisms behind technology usage, attempting to understand the impact of technology on society, to applied settings like security and knowing how dynamic behavioral patterns are. Further, by focusing on objective behavioral data over time (as seen across *Chapters II* to *IV*) as this captures users' interactions with specific technologies as they shift throughout the day and night. This is important, as discussed by Levine (2003), that when taking more traditional psychological approaches (e.g., experiments or interviews), participants are taking time out of their daily lives in order to participate. This then generates the data that the researcher attempts to develop into a coherent psychological narrative. However, in *Chapters II, III,* and *IV*, this issue is somewhat avoided as there has been arguably little to no impact on participants' daily life in order to understand their behavior as data was retrospectively collected in a naturalistic environment. Critically, this aims to provide more reliable and accurate ways to understand behavior and identity digitally.

By utilizing technology, we can better understand interactions with technology 'in situ', where we can utilize data from smartphones (via purpose built apps, for example: *Securacy* (Jones, Ferreira, Hosio, Goncalves, & Kostakos, 2015), *Apple Screen Time* (Ellis, Davidson, Shaw, & Geyer, 2019), apps (e.g., *PEG LOG* (Geyer, Ellis, & Piwek, 2018), *Funf in a Box* (Andrews, Ellis, Shaw, & Piwek, 2015).

Similarly, social scientists can scrape data from forums and other social media platforms, as seen in *Chapter III* and *IV*. These methods provide opportunities for social psychology to better understand real-world behavior. For instance, we can collect vast amounts of retrospective data or real-time data in order to understand behavioral patterns of usage to then potential infer 'identity' or social roles, predict future behavior, link digital identities, or find the online-offline behavior link.

These methods can be reliably used to understand interactions online, for instance, *Chapter III* demonstrated that users do indeed change roles over time. Additionally, we provided a way to understand these roles in terms of leadership and mapped out the most common role transitions users make, which has implications for marketing and advertisers (e.g., identification of new influencers) as well as for the identification of criminals within security settings (Katos & Bednar, 2008). However, the use of meta-data alone limits the inferences we can make about these groups of users. Hence, the use of content is another important approach to understanding user behavior. Of course, there are many methods analyzing content, from word counting (e.g., LIWC (Pennebaker, Booth, Boyd, & Francis, 2015) used in *Chapter IV*) to natural language processing techniques, pattern detection, or topic modelling (e.g., Verma et al., 2011). While the 'bag of words' approach to word counting provided insights concerning how users converged and diverged their linguistic style with a community (LSM), there are several other analyses that could be used to develop this research more (discussed in section 2). However, the analysis in *Chapter IV* did demonstrate that communication patterns of users is dynamic, much like user's reports of their self-presentation online seen in *Chapter V*. Hence, the two approaches in *Chapter III* and *IV* largely complement each other as one examines behavioral changes and the other examines communication changes of users. Similarly, they all demonstrate various shifts in behavior, which aligns with the findings in *Chapter V*.

In contrast, despite the shift away from behavioral measurement (Dolinski, 2018; Sassenberg & Ditrich, 2019), it is important to maintain understanding of user experiences with technology. Hence, *Chapter V* utilizes qualitative approaches to understand experiences with, attributes towards, and emotions about social media platforms and technologies. Of course, these findings are typically non-generalizable, however, the findings largely reflect the dynamic identity and self-presentation seen across the other empirical chapters of this thesis.

1.2 Ethical Implications

It is important to note that this new wave of data analytics has fueled an on-going and heated debate regarding data ethics, particularly in terms of behavioral analytics (UK Parliament, 2018). While there are indeed laws and regulations in place that aim to protect individuals, their identity, and privacy (e.g., GDPR); it is clear these are insufficient (Wachter & Mittelstadt, 2018). Effectively, while these laws may provide control to individuals about what data is collected, there is little to no control as to the inferences made about from this data (Wachter & Mittelstadt, 2018). Perhaps the most salient example is Cambridge Analytica, that created a widespread moral panic of digital privacy (or the lack of) (Lapowsky, 2019). Critically, this is an example of how not to ethically conduct large-scale data analytics, and academia can learn much from this in order to conduct high-quality and ethical research by harnessing opportunities various technologies affords us.

However, the data ethics debate remains critically important, and is unlikely to be 'solved' soon. While this debate moves quickly, Mittelstadt, Allo, Taddeo, Wachter, and Floridi (2016) discuss two overarching concerns regarding data ethics and inferences made from them: epistemic and normative. The former refers to problems with algorithmic decision-making producing merely probably knowledge, but this can be overly optimistic and unreliable in terms of inferences made (DeMasi, Kording, & Recht, 2017). This relates to the saying '*garbage in, garbage out*'. For instance, if the data quality is low or there is a lack of knowledge about the data, context, or understanding of what the data can actually reveal about those within it; the outputs will be arguably meaningless (DeMasi et al., 2017; Mittelstadt et al., 2016; Sawyer, 2019). The latter, normative concerns, typically relate to the 'fairness' of the outcomes of algorithms (Mittelstadt et al., 2016). The ways in which to handle this remain complex, however, via interdisciplinary work and accepting that social science input on 'big data' projects and developments is critically important in order to understand what can and cannot be inferred based on robust theory is also required.

With this in mind, the data and methods employed in this thesis are certainly what is known as behavioral analytics and great care was taken to anonymize users throughout this work. Additionally, inferences made remain largely exploratory and, we hope, fair to what the data was capable of demonstrating. The processes and methods used have been clearly stated with the hope to maintain transparency on how any conclusions were made about individuals and groups throughout thig thesis (and future work from it).

This thesis positioned itself as hoping to explore the opportunities of technology for psychological research (e.g., objective data sources, computational methods) as well as maintaining the perspective that technology is not inherently bad – nor good, where in all reality, much will have little to no effect on daily life (Orben & Przybylski, 2019). Conversely, technologies often change the way in which we make sense of the world around us (Major, Kaiser, O'Brien, & McCoy, 2007), or simply become ubiquitous leading to needing to engage with a technology as it is essential for everyday life. Answers typically lie between two extremes. Utilizing new technologies, theory, methods, and interdisciplinary collaborations will be critical to move researchers and societal discourse forward.

1.3 General Thesis Limitations

It is also important to consider limitations of this thesis, which can be broadly defined as theoretical or methodological. More specific limitations are addressed within each chapter.

From a theoretical standpoint, we could question the appropriateness of attempting to test theories that were developed prior to the internet – as these contexts differ in terms of being face-to-face and being able to 'hide' behind almost any identity one can imagine (Papacharissi, 2002). This could be taken a step further by stating that the field of 'cyberpsychology' as a whole has not developed new theory in line with changes to society via technological developments (Orben, 2018). This is problematic and as discussed in the Future Research section below, but does provide opportunities moving forward.

However, while considering the active and dynamic nature of behavior and identity, I did find evidence and support for social role theory online. Referring back to *Chapter II*, there remain questionable theoretical underpinnings of interactions with technology and naturally the implications of technology use. If we are to attempt to develop and understand the impacts of technology, there must be a shift towards balanced framing of research questions regarding technology and society (Davidson & Ellis, 2019).

In contrast, there remain challenges and limitations to all methods and approaches to research, however, using them in combination is helpful to mitigate these challenges. As noted in the ethical implications section, there are huge issues with data analytics and the use of machine learning techniques to infer behavior and to classify people (e.g., CV or mortgage screening). This thesis used methods and sampling techniques that were deemed appropriate for the research questions under investigation, which ranged from 'simple' statistics (*Chapter IV*: ANOVA, t-tests) to machine learning (*Chapter II* and *III*; K-MEANS and Naïve Bayes classifiers). These methods hoped to have been explained explicitly and clearly to show decision-making and the way the data was cleaned. Data for *Chapters II* and *III* has been shared, with code shared from *Chapter II*. *Chapter IV*'s data is readily available as well, which points towards Open Science and being transparent and reproducible. Additionally, pre- and post-prints are available for *Chapter II* and *Chapter III* being published with PLoS One means it is open access, which is ideal for disseminating research. Of course, other limitations will relate to the data itself. For instance, the findings from the ideological forums examined in *Chapter III* may not replicate if we used other forum data. However, this provides clear avenues for future research, for example: can we infer user leadership based on social roles across non-ideological communities? Here, we could draw from both *Chapters III* and *IV* and analyze simultaneously meta- and content-data to provide more in-depth analysis regarding user social roles online. Additional computational linguistics methods could be used as well, for instance, topic modelling and n-grams or bi-grams, which could reveal more about these groups.

2 Future Research Directions

In terms of future work relating to behavior change and adaptation across contexts, there are a number of future directions, which could extend, enhance, and develop from the research reported as part of this thesis.

2.1 'Basic' Interactions with Technology

As shown in *Chapters II, III,* and *IV*, exploratory research looking to understand what we can learn about user behavior based on meta-data or content data, is critically important for both theoretical and methodological development. For instance, in

Chapter III, we demonstrate that we can analyze changes in user behavior using the lens of user roles within a community. While much research has considered roles online (Welser et al., 2011; Welser, Gleave, Fisher, & Smith, 2007), we were the *first* to consider these roles over time and in terms of leadership within a community. We could consider additional metrics and data types (e.g., content) to examine roles further. Similarly, understanding at what point is a behavior change a change in role or perhaps a 'fluctuation' in behavior? There is also question regarding typical roles in online communities – some may have 'leaders' and a seemingly hierarchal system, however, does this always hold? This may include examining a variety of different forums, similar to Chan, Hayes, and Daly (2010), who examined role compositions across several online communities at a single point in time. Therefore, this could be extended by examining communities over time, examining individuals rather than groups of users, as well as potentially looking at the link between roles and communication style. For instance, Muir (2016) found that those conversing with someone in a superior 'role' is likely to exhibit greater levels of communication accommodation to encourage social approval, aligning with CAT theory (*Chapter IV*). Therefore, other work might want to consider whether there are more unique linguistic patterns or styles for specific types of roles in general, and whether these differences may be seen within individuals (e.g., the same person in different contexts) (e.g., Jones et al., 2012). Additionally, drawing from *Chapter IV*, there is potential to extend work focusing on the impacts of moderation online and whether this is an effective means to mitigate and reduce unruly behavior online.

Therefore, future work seeking to examine technology and society further may consider developing frameworks that are (unique) to online interaction. This may also inform mechanisms concerning *why* individuals continue to use a specific technology, which might include large-scale replications across various online contexts or devices. This could develop and extend theories to understand digital roles across online systems and devices, which would likely feed into current technology use theories (e.g., Technology Integration Model (Shaw, Ellis, & Ziegler, 2018)).

Contrastingly, it might include confirmatory work, which could consider how individuals change and adapt their behavior across time and contexts. Data collection will remain difficult; however, this could extend *Chapter IV* by comparing data from the same person from different devices (e.g., work phone versus personal phone), or various services (e.g., text messages, email, messenger, etc.) that could be simply textual or inclusion of profile scrapes and image. This could then be used to build models to test whether we can predict future behavior or if we can identify the same person in different contexts (e.g., Facebook and Twitter) or devices.

It is clear that we know comparatively little about *how* individuals and groups use various technologies, as discussed in *Chapters II* and *VI* in particular, and therefore, laying the foundations of technology usage studies is paramount (Van Rooij et al., 2018). Additionally, we must reevaluate our methods and analysis – including our data and the constructs it is indeed capturing (Flake & Fried, 2019). Embracing opportunities new technologies offer, from data sources, to new analytical capabilities, and visualization techniques is essential, alongside a movement towards interdisciplinary research. This therefore opens several avenues to consider psychometric measurement and development, which might be related to how we quantify behavior and what do these measurements reveal about individuals? Similarly, this may consider to what extent can we quantify 'identity'? While this research is taking place, in line with the prior ethical implications section, there is potentially substantial overlap in research between the social sciences and computer science/behavioral analytics moving forward.

2.2 Research Development: Finding the 'Missing Mechanism': Technology vs Psychology

When we consider engagement with technology, there appears to be a fairly polarized or binary view on the impacts of this technology use. New technologies develop rapidly, which poses a challenging environment for researchers to keep up with these new developments. Technology will continue to change and adapt the way we live (MacKenzie & Wajcman, 1999), which has also created a bizarre dualism with technology's integration into society, where on one hand, technology has, and continues to, transform our lives for the better (e.g., medicine, travel, communication developments). Yet, on the other hand, there is a consistent and powerful voice within the social sciences in particular, that pathologizes new technologies. Interestingly, other disciplines more commonly adopt a view that technologies (e.g., wearables) can greatly improve and rejuvenate public health (e.g., Fisch, Chung, & Accordino, 2016; Munos et al., 2016).

This binary logic that technology is either good or bad, needs to become more nuanced to ensure we do not miss the bigger picture. People are not binary; therefore, research must view them accordingly as flexible, analogue beings in a digital world and be mindful of the wider implications of technology and society (e.g., Hassan, 2008; Martin, 2008). Hence, research that focuses on the implications of engagement of technology must find or develop suitable theory to underpin this research (e.g., Orben, 2018), which perhaps ought to consider research from wider perspectives than psychology, e.g., Science and Technology Studies (STS) and Information Systems (IS). Currently, social psychological work tends to utilize theories or frameworks that are unconvincing (e.g., addictions frameworks or social learning theory/cognitive-behaviorist models), which continue to lack sufficient evidence for continued use (Starcevic, Billieux, & Schimmenti, 2018). There appears to be a 'missing mechanism' in terms of 'basic' interactions with technology, the influence of technology on society, and whether this is a bi-directional influence.

Technologies are hugely interwoven into society, where we have become reliant on technologies from transport, communications, work, warfare, to science – this interaction between people, society, and technology is inherently complex (MacKenzie & Wajcman, 1999), which has indeed been long-discussed across STS and IS, as noted in *Chapter I*. This means social psychology has a well-documented foundation to potentially build on moving forward (e.g., Actor-Network Theory, Technological Determinism, etc.). This perhaps will better our understanding of how and why people use technology prior to attempting to predict and examine the impacts, implications, and in some cases – behavioral or other interventions via technology. Once there are more robust theories or mechanisms understanding technology usage, engagement, and interaction – this will naturally lead and underpin future applied research (e.g., health, security, marketing) based on stronger foundations.

While attempting to bridge the gap between online and offline behavior remains difficult, (new) technologies again have much to offer this area of research (e.g., Sapiezynski, Stopczynski, Wind, Leskovec, & Lehmann (2018) used a custom built Android app to track location, usage, and specific app usage in order to analyze the online and offline behavior of students). The ability to predict offline behavior or to even link a person to various online accounts is challenging, however, greater ability to do so is important for a variety of contexts (e.g., security, health (e.g., Lyons et al., 2009)). From a security perspective, this might include studies attempting to test whether those who state they are ‘going’ to an event on Facebook or Twitter actually attend these events. Further, with more in-depth data analysis, one might be able to find identifiers that can predict likelihoods of different people’s offline behavior. Understanding this link, if there is a reliable one, would be of huge benefit to better handling ‘shitposters’ or ‘trolls’ online, and whether these users may actually act upon actions posted online. However, gathering data for such studies remains difficult and reliant on a mixture of secondary data analysis (e.g., scraping data) and perhaps speaking directly to participants to confirm their attendance, or the use of (smartphone) apps (e.g., location tracking).

3 Final Comment

Technology is encompassing and persuasive. During the course of this thesis, it has become apparent to me that many disciplines are, in one way or another, interested in the interaction between technology, people, and society. This has inevitably led to disagreements within and across these disciplines (whether that is due to approaches, methods, or philosophical standpoints). A surprising amount of research remains siloed, which while not a new perspective in itself, appears rather ironic in the context of technology that has provided so many new avenues to share and collaborate. This thesis has attempted to bridge some of these gaps by noting and utilizing opportunities from overlaps between social psychology, computer science, data science, communications, and even touching on STS both here and in *Chapter I.* It is critically important for the social sciences to join with others in an interdisciplinary context. There is much to learn from various perspectives, which will enrich future research. The whole will then be greater than the sum of its parts.

References

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS ONE*, *10*(10), 1–9. https://doi.org/10.1371/journal.pone.0139004

Chan, J., Hayes, C., & Daly, E. M. (2010). Decomposing Discussion Forums using User Roles. *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*, *112*(988), 215–218. Retrieved from http://aran.library.nuigalway.ie/xmlui/handle/10379/1114

Davidson, B. I., & Ellis, D. A. (2019). Social media addiction: technological déjà vu. *BMJ*, *365*(I4277). https://doi.org/10.1136/bmj.l4277

DeMasi, O., Kording, K., & Recht, B. (2017). Meaningless comparisons lead to false optimism in medical machine learning. *PLoS ONE*, *12*(9), e0184604. https://doi.org/10.1371/journal.pone.0184604

Dolinski, D. (2018). Is psychology still a science of behaviour? *Social Psychological Bulletin*, *13*(2), e25025.

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies: Special Issue on Human Accuracy*. https://doi.org/10.1016/j.ijhcs.2019.05.004

Erikson, E. (1959). *Identity and the Life Cycle*. London, UK: W. W. Norton & Company Ltd.

Fisch, M. J., Chung, A. E., & Accordino, M. K. (2016). Using technology to improve cancer care: social media, wearables, and electronic health records. *American Society of Clinical Oncology Educational Book*, *36*, 200–208.

Fiske, S. T. (2010). *Social Beings* (2nd ed.). NJ, USA: Wiley.

Flake, J. K., & Fried, E. I. (2019). Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. *PsyArXiv*. https://doi.org/10.31234/osf.io/hs7wm

Geyer, K., Ellis, D. A., & Piwek, L. (2018). A simple location-tracking app for psychological research. *Behaviour Research Methods*. https://doi.org/https://doi.org/10.3758/s13428-018-1164-y

Hassan, R. (2008). *The Information Society*. Cambridge, UK: Polity Press.

Jones, S., Cotterill, R., Dewdney, N., Muir, K., & Joinson, A. (2012). Finding Zelig in Text: A Measure for Normalising Linguistic Accommodation. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers* (pp. 455–465).

Jones, S. L., Ferreira, D., Hosio, S., Goncalves, J., & Kostakos, V. (2015). Revisitation analysis of smartphone app use. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1197–1208). Osaka, Japan. https://doi.org/10.1145/2750858.2807542

Katos, V., & Bednar, P. M. (2008). A cyber-crime investigation framework. *Computer Standards and Interfaces*, *30*, 223–228.

Lapowsky, I. (2019). How Cambridge Analytica sparked the great privacy awakening. Retrieved May 23, 2019, from https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/

Levine, M. (2003). Times, Theories and Practices in Social psychology. *Theory & Psychology*, *13*(1), 53–72.

Lyons, R. A., Jones, K. H., John, G., Brooks, C. J., Verplancke, J.-P., Ford, D. V, … Leake, K. (2009). The SAIL databank: linking multiple health and social care datasets. *BMC Medical Informatics and Decision Making*, *9*(3). https://doi.org/10.1186/1472-6947-9-3

MacKenzie, D., & Wajcman, J. (1999). *The social shaping of technology* (2nd Editio). Open University Press.

Major, B., Kaiser, C. R., O’Brien, L. T., & McCoy, S. K. (2007). Perceived Discrimination as Worldview Threat or Worldview Confirmation: Implications for Self-Esteem. *Journal of Personality and Social Psychology*, *92*(6), 1068–1086. https://doi.org/10.1037/0022-3514.92.6.1068

Martin, A. (2008). Digital Literacy and the “Digital Society.” In *Digital literacies: Concepts, politics and practices* (pp. 151–176). Peter Lang Inc.

Matias, J. N. (2019). Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *Proceedings of the National Academy of Sciences*, 201813486. https://doi.org/10.1073/pnas.1813486116

Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Muir, K. (2016). Power and linguistic accommodation in cmc. In *In the proceedings of the 66th International Communication Association Annual Conference*. Fukuoka, Japan.

Munos, B., Baker, P. C., Bot, B. M., Crouthamel, M., Vries, G. de, Ferguson, I., …

Wang, P. (2016). Mobile health: the power of wearables, sensors, and apps to transform clinical trials. *Annals of the New York Academy of Sciences*, *1375*(1), 3–18.

Orben, A. (2018). Cyberpsychology: A field lacking theoretical foundations. *PsyPAG*, (107), 12–14.

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, *3*, 173–182. https://doi.org/10.1038/s41562-018-0506-1

Papacharissi, Z. (2002). Presentation of Self In Virtual Life: Characteristics of Personal Home Pages. *Journalism & Mass Communication Quarterly*, 643–660.

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic inquiry word count: LIWC2015. Pennebaker Conglomerates*. Austin, Texas, USA. Retrieved from www.LIWC.net

Preece, J., & Schneiderman, B. (2009). The Reader-to-Leader Framework: Motiviating Technology-Mediated Social Participation. *Transactions on Human-Computer Interaction*, *1*(1), 1–25. https://doi.org/10.5121/ijfcst.2014.4403

Sapiezynski, P., Stopczynski, A., Wind, D. K., Leskovec, J., & Lehmann, S. (2018). Offline Behaviors of Onine Friends. *ArXiv Preprint ArXiv:1811.03153v2*.

Sassenberg, K., & Ditrich, L. (2019). Research in Social Psychology Changed Between 2011 and 2016: Larger Sample Sizes, More Self-Report Measures, and More Online Studies. *Advances in Methods and Practices in Psychological Science*, *2515245919*.

Sawyer, J. (2019). What Social Word Got Right and Why It's Needed for our [Technology] Revolution. Retrieved June 5, 2019, from https://jods.mitpress.mit.edu/pub/sawyer

Shaw, H., Ellis, D. A., & Ziegler, F. V. (2018). The Technology Integration Model (TIM). Predicting the continued use of technology. *Computers in Human Behavior*, *83*, 204–214. https://doi.org/10.1016/j.chb.2018.02.001

Starcevic, V., Billieux, J., & Schimmenti, A. (2018). Selfitis and behavioural addiction: A plea for terminological and conceptual rigour. *Australian & New Zealand Journal of Psychiatry*, *52*(10), 919–920.

UK Parliament. (2018). Algorithms in decision-making. Retrieved May 23, 2019, from
https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pd

f

Van Rooij, A. J., Ferguson, C. J., Colder Carras, M., Kardefelt-Winther, D., Shi, J., Aarseth, E., ... Przybylski, A. K. (2018). A weak scientific basis for gaming disorder: Let us err on the side of caution. *Journal of Behavioral Addictions*, *7*(1), 1–9.

van Zoonen, L. (2013). From identity to identification: Fixating the fragmented self. *Media, Culture and Society*, *35*(1), 44–51. https://doi.org/10.1177/0163443712464557

Verma, S., Vieweg, S., Corvey, W. J., Palen, L., Martin, J. H., Palmer, M., … Anderson, K. M. (2011). Natural Langauge Processing to the Rescue? Extracting “Situational Awareness” Tweets During Mass Emergency. In *In the proceedings of the fifth Interational AAAI Conference on Weblogs and Social Media* (pp. 385–392).

Wachter, S., & Mittelstadt, B. (2018). A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI [forthcoming]. *Columbia Business Law Review*, 1–84.

Welser, H., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). Finding social roles in Wikipedia. *Proceedings of the 2011 IConference on - IConference ’11*, 122–129. https://doi.org/10.1145/1940761.1940778

Welser, H. T., Gleave, E., Fisher, D., & Smith, M. A. (2007). Visualizing the Signatures of Social Roles in Online Discussion Groups. *Journal of Social Structure*, *8*(2), 1–32.